Journal of the International Neuropsychological Society (2007), **13**, 626–634.
Copyright © 2007 INS. Published by Cambridge University Press. Printed in the USA.
DOI: 10.1017/S1355617707070786

The random number generation task: Psychometric properties and normative data of an executive function task in a mixed sample

MAARTEN PETERS, TIMO GIESBRECHT, MARKO JELICIC, AND HARALD MERCKELBACH
Department of Experimental Psychology, Maastricht University, Maastricht, The Netherlands

(RECEIVED June 29, 2006; FINAL REVISION February 2, 2007; ACCEPTED February 2, 2007)

Abstract

We investigated the psychometric properties of the random number generation (RNG) task in four studies using a mixed sample of young adults ($n = 306$), middle-aged adults ($n = 40$), and patients diagnosed with schizophrenia ($n = 26$). Data in study 1 were best accounted for by a three-factor solution representing inhibition of stereotypical schemas (seriation), output inhibition (repetition), and monitoring of previous output (cycling). Modest test-retest correlations were found, with the seriation factor showing acceptable stability across time (study 2). In study 3, RNG task performance was related to scores on concurrent neurocognitive tasks to establish construct validity. RNG scores correlated with healthy controls' performance on the Stroop color-word test and patients diagnosed with schizophrenia with executive dysfunctions. Patients diagnosed with schizophrenia performed poorer on the seriation factor of the RNG than healthy control participants (study 4). Our results indicate that the RNG task has modest to acceptable psychometric properties. It primarily taps executive subfunctions (i.e., inhibition, updating, and monitoring), which are affected by psychopathological or neurological deficits. (*JINS*, 2007, *13*, 626–634.)

Keywords: Factor analysis, Psychometrics, Inhibition, Schizophrenia, Reliability, Validity

INTRODUCTION

When asked to generate random sequences of digits, people usually perform poorly (i.e., non-random; Wagenaar, 1972). A paradigm to measure deviations from randomness is the Random Number Generation (RNG; e.g., Ginsburg & Karpiuk, 1994) task. In this task, participants are asked to produce sequences of digits (e.g., 1–10) in a random fashion. Successful RNG performance requires various higher order processes, including retaining task-related instructions (e.g., set size, task instructions) in memory, integrating information, and holding it "on-line" in working memory (central executive involvement; Baddeley, 1986), avoiding interference, monitoring output, and switching or modifying production strategy in accordance with the "on-line" concept of randomness (executive functioning; Baddeley et al., 1998; Jahanshahi et al., 2006). There is convincing evidence that people's difficulties with RNG are neither attributable to a misconception of randomness nor to short-term memory problems (Baddeley, 1998; Wagenaar, 1970).

Several versions of RNG have been used (e.g., Brugger, 1997). They differ in set size (0–9, 1–20, etc.; e.g., Towse, 1998), pacing technique (paced or unpaced; e.g., Joppich et al., 2004), response pace (500 msec, 1 sec, etc.; e.g., Daniels et al., 2003), response modality (oral, written, etc.; e.g., Schneider et al., 2004), and instructions used (implicit, explicit, biased; see for a review Brugger, 1997). Despite these differences, there is broad consensus that RNG requires the allocation of central executive resources (e.g., Baddeley, 1986).

Several RNG parameters have been proposed to quantify deviations from randomness (e.g., Ginsburg & Karpiuk, 1994; 1995; Towse & Neil, 1998). One influential set of RNG parameters is described by Ginsburg and Karpiuk (1994). It consists of the following 9 parameters: Coupon (Cn), Gap (Gp), Poker (Pk), Runs (Rn), Repetitions (Rp), Series (Sr), Variance of digits (VD), Digram repetition (DR), and Cluster ratio (Cr). Table 1 gives definitions of these 9 parameters. In their study, Ginsburg and Karpiuk (1994) had 32 undergraduates (3 men), ranging in age from 19 to

Correspondence and reprint requests to: Maarten Peters, Department of Experimental Psychology, Maastricht University, PO Box 616, 6200 MD, Maastricht, The Netherlands. E-mail: m.peters@psychology.unimaas.nl

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

Table 1. Eigen values, percentage variance explained, and factor loadings of the RNG three-factor solution for undergraduate students (n = 306). Definitions of the different RNG indices are also shown.

	Definition	Factor 1 Seriation	Factor 2 Cycling	Factor 3 Repetition
Eigenvalue		2.94	2.09	1.44
% Variance explained		32.61	23.17	15.97
Factor loadings				
Cluster Ratio	Variance of 100 successive responses in the digram matrix, which presents the frequency with which each item is followed by each of the possible other items. In a series with n there are n^2 possible parings.	.871	—	—
Digram Repetition	Number of digram repetitions (e.g., 1,5, . . . 1,5).	.850	—	—
Series	Number of consecutive digrams (e.g., 4,5). Score is calculated by summing up the number of these digrams.	.803	—	—
Runs	Number of responses in successive ascending runs. The runs score is the variance of these numbers.	.756	—	—
Gap	Measure of cycling through the set of 10 items. This measure is obtained by counting the number of gaps between two identical digits. The median of this number is then calculated.	—	−.849	—
Variance of digits	Variance of the frequencies of each digit.	—	.831	—
Coupon	Number of responses that occur before all digits are emitted. Coupon score is the average for all complete sets.	—	.419	—
Repetition	Number of identical pairs (e.g., 4,4). Score is the sum of the number of these identical pairs.	—	—	.932
Poker	Number of repetitions of the same digit within the 20 sequences of 5 successive responses. The sum of the number of sequences that contain exactly two of a kind is the Poker score.	—	—	.903

Note. Factor loadings <.40 are omitted in this table

50 years (M = 30), produce a sequence of 100 digits consisting of the digits 0 to 9, whereas avoiding any system. RNG was paced by a metronome at 40 responses/min. Next, the authors performed a factor analysis (Principal Component Analysis, Varimax rotation) on the described RNG indices. This yielded three factors: seriation (loadings Rn = .84, DR = .79, and Sr = .77), cycling (loadings Gp = .86, VD = .81, Cn = .66, and Pk = .50), and repetition (loadings Rp = .91, Pk = .78, and Cn = .53). The three factors are interpreted as reflecting inhibition of stereotyped cognitive schemas, successful monitoring of previous output, and output inhibition, respectively (Williams et al., 2002).

Although RNG has been widely used as a research tool in healthy and clinical populations (Artiges et al., 2000; Brown et al., 1998; Brugger et al., 1996; Joppich et al., 2004), psychometric data (e.g., factor structure, test-retest reliability, construct validity) about this tool are scarce. With this in mind, we conducted four studies to investigate the psychometric properties of the RNG task, focusing on factor structure of the indices proposed by Ginsburg and Karpiuk (1994) (study 1), test-retest reliability and practice effects (study 2), construct validity (study 3), and criterion-related validity (study 4) in a mixed sample of healthy participants and clinical patients.

METHODS

Study 1: Factor Structure

The three-factor solution proposed by Ginsburg and Karpiuk (1994) was based on a small sample (n = 32). Because their three-factor solution is generally in accordance with the taxonomy of response biases in human behavior (Rabinowitz, 1970), we wanted to examine whether we could replicate the Ginsburg and Karpiuk (1994) solution, now using a more appropriate sample size for conducting factor analysis.

Participants

A group of 306 (98 men) undergraduate psychology students participated in this study in return for course credits. Age ranged from 17–54, with a mean age of 19.90 (SD = 4.37). None of the participants had a history of alcoholism, head injury, psychiatric illness, or a neurological condition. The study was approved by the standing ethical committee of the Faculty of Psychology, Maastricht University. Note that the data described in this manuscript were obtained in compliance with the regulations of our institution, and human

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

research was completed in accordance with the Helsinki Declaration (http://www.wma.net/e/policy/b3.htm).

Materials and Procedure

Participants were tested individually. Upon arrival in the laboratory, they signed an informed consent form and were administered the RNG task. The RNG task was taken from Towse (1998), with the exception of response pace, which was set at one digit per sec (indicated by a metronome adjusted to 60 bpm). This was done to increase comparability with other factor analytic studies (e.g., Miyake et al., 2000) and studies relying on similar samples (e.g., Brugger et al., 1995). More specifically, participants were asked to generate a random sequence of digits (set size: 1–10), for a period of 100 sec. The concept of randomness was explained using the instruction of Baddeley (1966), which draws an analogy of picking digits out of a hat, reading them loud, putting them back and then picking the next digits from the hat (see also Towse, 1998). Our instruction emphasized that a random sequence would not contain a preponderance of repetitions or adjacent number values.[1] Participants were told to say out loud one digit with each tone. They were told that if their response rate fell behind the required pace, they were to listen to the rhythm and speed up their response time accordingly.

Data Analysis

The 9 RNG indices (Ginsburg & Karpiuk, 1994), including cluster ration (CR), were calculated (cf. supra). These 9 indices were subjected to Principal Component Analysis (PCA) with an orthogonal (varimax) as well as an oblique rotation (direct oblimin), because we did not know whether the extracted factors would correlate with each other (oblique) or not (orthogonal).

Our selection of factors was based on both a scree plot of eigen values and Kaiser's criterion (Kaiser, 1960) with the cut-off point set at 1. Furthermore, only factor loadings greater than .4 were considered (Stevens, 1992). Of course, theoretical meaningfulness of the resulting factor structure was also taken into account.

RESULTS

The PCA yielded three factors with eigen values greater than 1 (see Table 1). After rotation, both orthogonal and oblique rotations yielded a similar factor structure. For this reason, the results from the simpler orthogonal (varimax) rotation are presented here.[2] The factor solution basically replicates that of Ginsburg and Karpiuk (1994; see Table 1). However, in contrast to the factor solution reported by Ginsburg and Karpiuk, we found that Cn did not load on the repetition factor, whereas Pk exclusively contributed to the repetition factor.

DISCUSSION

Together with those of Ginsburg and Karpiuk (1994), our findings imply that the 9 RNG indices can be grouped into three clusters. The fact that these three factors represent orthogonal dimensions suggests that they tap different aspects of executive functioning. Repetition consists of the rehearsal of the same digit in succession, with excessive repetition being related to general deficits in suppression of previous responses (i.e., output inhibition; Bradshaw & Mattingley, 1995). Seriation can best be understood as an inability to suppress stereotypical schemas (e.g., Williams et al., 2002), like counting forward, backward, by two's and so forth. This bias can be interpreted as the consequence of interference by overlearned tendencies to arrange numbers according to their natural order. Cycling occurs when individuals attempt to systematically use every possible alternative before repeating any digit, which means that they successfully monitor previous output (e.g., Williams et al., 2002).

METHODS

Study 2: Test-retest Stability

To investigate temporal stability of the RNG indices, the task was administered twice to a subsample of healthy controls and patients diagnosed with schizophrenia with an interval of two weeks. We hypothesized that the RNG scale would show satisfactory stability.

Participants

Participants were 59 young adults (subsample of study 1; 17 men) and 10 (8 men) of a total of 26 inpatients diagnosed with schizophrenia (see studies 3 and 4). Mean age was 19.27 years ($SD = 1.54$; Range = 18–26) for the young adult sample and 37.40 years ($SD = 11.81$; Range = 18–59) for the 10 patients. Mean educational level of the patients was 4.80 ($SD = 1.03$; anchors: 1 = *lower education*; 7 = *university degree*; Verhage, 1964). Patients diagnosed with schizophrenia were recruited from two psychiatric hospitals in Belgium. Diagnoses were based on DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, 4th edition; American Psychiatric Association, 1994) criteria

[1] One could speculate that these instructions may influence the RNG outcome measures of repetition avoidance or serial responding. However, several studies (Peters et al., 2006; Towse, 1998) have found that healthy participants who have received these instructions commit qualitatively and quantitatively similar errors as those without such warning (e.g., Giesbrecht et al., 2004; Ginsburg & Karpiuk, 1994).

[2] When looking at the underlying correlations between the extracted factors in the direct oblimin PCA, no significant correlations were apparent (all $r's < .15$). This shows that the three extracted factors are independent, thereby supporting the use of Varimax PCA.

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

Table 2. Test-retest reliability and practice effects of the three factor scores (z-scores) for controls ($n = 59$) and patients diagnosed with schizophrenia ($n = 10$). Means (*SD*) are given for T1 and T2.

Group	RNG factors	Test-retest	*M* (*SD*) T1	*M* (*SD*) T2	*t* (58)
		Pearson			
Controls	Seriation factor	.51**	−.23 (.43)	−.22 (.53)	−.01, n.s.
$n = 59$	Cycling factor	.18	−.05 (.56)	.02 (.48)	−.62, n.s.
	Repetitions factor §	.42**	−.07 (.73)	.05 (.97)	−1.17, n.s.
		Spearman			*Z* statistic #
Schizoprenia	Seriation factor	.43	1.35 (1.59)	1.25 (1.31)	−.15, n.s.
$n = 10$	Cycling factor	.67*	.28 (.40)	−.11 (.34)	−.18, n.s.
	Repetitions factor	.20	.43 (1.50)	−.31 (.42)	−.38, n.s.

* = $p \leq .05$
** = $p < .001$
§ = Spearman's ρ correlations
# Wilcoxon's signed rank test; (Skewness repetition = 2.27)

for schizophrenia and were made by a team of experienced psychiatrists who conducted structured diagnostic interviews. All patients were on fixed doses of antipsychotic medication, either typical (88%) or atypical (12%).[3] Duration of illness of the schizophrenic sample was 5.81 years (*SD* = 6.21). None of the participants had a history of severe neurological disorders, substance abuse, or (co-morbid) DSM-IV axis 1 or 2 disorder. The study was approved by the standing ethical committee of the Faculty of Psychology, Maastricht University.

Materials and Procedure

Materials and procedures in session one (RNG1) were identical to those in study 1. During session two (two weeks later), these samples had the RNG task administered for a second time (RNG2).

Statistical Analysis

Using an α of .05, two-tailed, test-retest stability (using Pearson and Spearman correlations) and practice effects (paired samples *t*-tests and signed rank test) were explored using the three RNG factors established in study 1.[4] To this end, separate RNG parameters were z-transformed and factor scores were computed by averaging across relevant RNG indices.

RESULTS

Test-retest stability and practice effect data are summarized in Table 2. In the subsample of healthy controls, we found the highest test-retest correlation for the seriation factor. In the schizophrenic subsample ($n = 10$), the highest test-retest correlation (Spearman's ρ) was found for the RNG factor cycling. In the healthy as well as the schizophrenic sample, no practice effects were found for the three factor scores.

DISCUSSION

In healthy controls, RNG factors seem to possess at best modest test-retest stability. Meanwhile, with repeated administration, healthy controls did not show significant practice effects for the three factors. For the schizophrenic sample, highest test-retest stability was found for the cycling factor, with no practice effects on the three factors. Test-retest correlations of the RNG scales in healthy controls and our clinical sample failed to reach the minimum of .80 required for a clinical psychometric instrument (Anastasi & Urbina, 1997; see also De Zubicaray et al., 1998; Jelicic et al., 2001).

METHOD

Study 3: Construct Validity

In this study, we investigated whether the RNG factors seriation, repetition, and cycling are related to specific neurocognitive tasks that are known to tap the constructs of inhibition of stereotypical schemas, output inhibition, and monitoring of previous output. Firstly, based on previous research (Brugger et al., 1995), we hypothesized that a failure to inhibit stereotypical schemas (i.e., heightened seriation) would positively correlate with interference susceptibility measured by the Stroop task (Stroop, 1935). Secondly, because keeping and updating information "on-line" is important for accurate monitoring of previous output and output inhibition, we expected a relationship between the central executive "online" component of working memory (backward digit span; Gerton et al., 2004), and the repeti-

[3]Previous research has found that randomization performance in schizophrenia may improve with onset of neuroleptic medication due to an improvement of concentration, but soon declines again to off-medication baseline (e.g., Axmacher et al., 1970).

[4]Here, we describe the standardized factor scores for the test-retest stability and construct validity of the RNG. Results for separate RNG indices can be obtained from the first author.

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

tion and cycling factors. Finally, we hypothesized that the RNG factors would relate to more unitary executive function tasks in a clinical sample of patients diagnosed with schizophrenia (see for example Miyake et al., 2000).

A typical finding in RNG studies is that when processing demands increase (e.g., faster response pace), deviations from randomness also become more marked (e.g., Jahanshahi et al., 2006; Wagenaar, 1970). We sought to explore whether individual differences in processing speed would show a similar linear relationship with deviations from randomness. Furthermore, it has been argued that RNG is not purely driven by a limitation in non-executive working memory span (e.g., Baddeley, 1966; Wagenaar, 1970). We wanted to directly test this by relating the RNG factors to individual differences in non-executive working memory (forward digit span).

Participants

This study involved a schizophrenic subsample ($n = 26$; 21 men) and a young adult subsample (see study 2; $n = 59$). Mean age for the schizophrenic subsample was 36.35 years ($SD = 12.83$; Range = 18–71). Mean educational level of the schizophrenic subsample was 4.54 ($SD = 1.39$). Duration of illness (in years) was 6.52 ($SD = 7.21$).

Materials and Procedure

Apart from the RNG task, the young adult subsample was administered the forward and backward digit span task and the Stroop color-word test. In the schizophrenic subsample, the Behavioral Assessment of Dysexecutive Syndrome (BADS; Wilson et al., 1998) and the Wisconsin Card Sorting Test (WCST; Heaton et al., 1993) were administered. Patients diagnosed with schizophrenia were tested during standard neuropsychological screening protocols. For this reason, we did not have the opportunity to also collect the digit span and Stroop color-word test in this sample. For the healthy sample, WCST and BADS were not administered, because it is known that these instruments were designed to assess executive functions in clinical populations. Thus, these measures usually yield ceiling effects in (normal) healthy controls.

Digit Span

The forward and backward digit span tests (for a full description see Stinissen et al., 1970) were administered. Each subtest was stopped after two subsequent incorrect reproductions. The number of correct orally produced strings in each subset was used as outcome measure.

Stroop Color-word Test

The classic Stroop color-word test (Stroop, 1935) was used in which participants are asked to read aloud or name the stimuli on each card (color names of card 1, color of the patches on card 2, and color of the ink on card 3) one after the other as quickly as possible but without making errors. Correcting errors was allowed. However, given the infrequency of errors in this sample (mean error score $<.50$), they were discarded in further analyses. As an index of processing speed, time to read card 2 (T2) was measured. Susceptibility to interference was calculated by subtracting T2 from the time needed to name the colors of the ink of card 3 (T3).

Behavioral Assessment of the Dysexecutive Syndrome (BADS)

This task comprises 6 subtasks (see for a Dutch version, Krabbendam & Kalff, 1998) and is used as a measure of executive functioning. In the current study, total profile scores (maximum = 24) were used, with higher scores indicating better executive functioning.

Wisconsin Card Sorting Test (WCST)

A computerized version of the WCST was administered (128 test trials; for a full description see Heaton et al., 1993). For the present analyses, the WCST parameters "categories completed" (0–6), and "number of perseverative errors" were extracted.

RESULTS

Table 3 shows how RNG factors relate to neurocognitive tasks. For healthy controls, a high color naming speed was associated with a heightened seriation score. Also, a modest positive correlation was found between the RNG seriation factor and the Stroop-interference measure. For the repetition factor, a modest but significant and negative correlation was found with forward digit span. All other correlations remained non-significant. In the schizophrenia sample, the RNG factor scores of seriation and cycling correlated negatively with the BADS total score.

DISCUSSION

In this study, we made an attempt to relate RNG factors to various neurocognitive tasks. Significant correlations were primarily found for the RNG seriation factor, albeit that these correlations were modest. Also, with so many correlations, there is the risk of experimenter-wise errors. On the other hand, the significant correlations that did emerge are theoretically meaningful. For example, RNG seriation correlated positively with Stroop interference, which is not surprising when one considers that RNG seriation reflects difficulties in inhibiting stereotype responses. In this respect, our findings come close to those of Brugger et al. (1995), who reported a modest correlation ($r = .30$) between Stroop interference and counting bias. We also found that a high response speed (as indexed by Stroop color naming) is related

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

Table 3. Correlations of RNG factors with Stroop color naming, Stroop-interference, and digit span (forward and backward) in healthy controls ($n = 59$) and with Behavior Assessment of Dysexecutive Syndrome (BADS) and Wisconsin Card Sorting Test (WCST) in patients diagnosed with schizophrenia ($n = 26$).

Group	Task	Seriation factor	Cycling factor	Repetition factor
Controls	Stroop color naming	.41**	−.08	−.11
$n = 59$	Stroop interference	.30*	−.02	.07
	Digit span Forward	−.24	−.03	−.29*
	Digit span Backward	−.20	−.08	−.11
Schizophrenia	WCST number trials	−.04	−.27	−.03
$n = 26$	WCST pers. errors	.07	.32	.26
	BADS total score	−.54**	−.48**	−.32

*$p < 0.05$
**$p < 0.01$

to heightened seriation, which is not surprising if one assumes that failure to inhibit stereotypes is a trade-off of high response speed. In the schizophrenic subsample, we found significant negative correlations between RNG factors seriation and cycling and BADS scores, which is a first indication that these RNG factors are related to a more unitary executive functioning task (see Miyake et al., 2000). The significances (i.e., $p < .01$) of these correlations were such that they would survive Bonferroni corrections for multiple testing.

Study 4: Criterion-related Validity

Several studies investigating RNG deficits in schizophrenia (e.g. Artiges et al., 2000; Horne et al., 1982; Rosenberg et al., 1990; Salamé et al., 1998; Shinba et al., 2000), noted that patients diagnosed with schizophrenia have an increased tendency to produce stereotyped series and repetitive responses. Using the Ginsburg and Karpiuk factors (1994), we made an attempt to replicate this pattern. More specifically, we hypothesized that patients diagnosed with schizophrenia would show more extreme scores on the RNG seriation and repetition factors compared to healthy control participants (young and mid-age).

Apart from psychopathology as possible predictor of RNG performance, there is the issue of aging. Van der Linden et al. (1998) were the first to find that elderly participants (age range 60–70) produce on random generation tasks more series but not more repetitions in comparison to young adults (age range 20–30). This is probably because of the demands that such tasks place on the central executive capacity of the elderly participants. We were interested whether a similar age-related decline in seriation factor would be found in a middle-aged group (aged 40–60) in comparison to young adults.

METHOD

Participants

In this study, data of studies 1–3 were collapsed and further extended with a middle-age subsample. Thus, study 4 relied on the schizophrenic subsample ($n = 26$), the young adult subsample ($n = 299$; now with specific age range 18–25), and a middle-aged subsample ($n = 40$; 17 men; Mean age= 48.14; $SD = 8.56$; age range 40–62; hereafter mid-age). Mean educational level of the mid-age sample was 5.03 ($SD = 1.05$)

Materials and Procedures

Materials and procedures were identical to those used in study 1.

RESULTS

One-way Analyses of Variance (ANOVAs) were carried out for the separate RNG factors. Apart from the mean z-transformed factor scores, mean scores on the 9 different RNG indices are also given in Table 4 for normative purposes. The only effect was a significant effect of group status on the seriation factor.[5] A Games-Howell post-hoc test revealed that patients diagnosed with schizophrenia scored higher than young adult and the mid-age subsample ($p < .01$).[6] No significant differences emerged between the young adult and mid-age subsample (all p's $> .05$).

[5]Separate one-way ANOVA's were also carried out with groups being the Study 2 young adult subsample ($n = 59$), mid-age subsample ($n = 40$), and the schizophrenic subsample ($n = 26$), using post-hoc Bonferroni corrections. These analyses yielded similar results.

[6]The Games-Howell post-hoc procedure is designed to analyze data from unbalanced designs in which sample variances differ (e.g., Field, 2005).

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

Table 4. Normative data (means and *SD*) of the RNG factors (z-transformed factor scores) and individual RNG indices (absolute scores) for young ($n = 299$) and mid-age ($n = 40$) healthy controls, and the schizophrenic sample ($n = 26$).

	Young Mean (*SD*)	Mid-age Mean (*SD*)	Schizophrenia Mean (*SD*)	Statistics $F(2,362)$
Factor				
Seriation	−.08 (.69)	−.03 (.71)	.98 (1.73)	>15.0, $p < .001$
Cycling	.00 (.43)	.05 (.57)	−.03 (.34)	n.s.
Repetition	−.01 (.96)	.08 (.75)	−.06 (.77)	n.s.
RNG indices				
Cn	20.87 (8.39)	23.21 (12.26)	18.58 (6.26)	n.s.
Gp	8.36 (.68)	8.31 (.65)	8.33 (.96)	n.s.
Pk	4.35 (3.27)	4.60 (2.73)	4.15 (3.27)	n.s.
Rn	1.17 (.50)	1.17 (.49)	2.06 (1.92)	>6.0, $p < .005$
Rp	1.40 (2.83)	1.76 (2.42)	1.27 (1.80)	n.s.
Sr	33.62 (8.27)	34.17 (8.49)	41.08 (14.32)	>6.0, $p < .005$
VD	7.62 (5.25)	7.31 (4.39)	8.79 (5.85)	n.s.
DR	43.37 (4.55)	43.95 (4.69)	48.96 (8.41)	>6.0, $p < .005$
Cr	1.41 (.30)	1.44 (.33)	1.83 (.68)	>6.0, $p < .005$

DISCUSSION

Criterion-based validity of the RNG task is most promising for the seriation factor, because this factor was able to differentiate between patients diagnosed with schizophrenia and healthy controls. This accords well with previous studies showing a strong counting bias in patients diagnosed with schizophrenia (e.g., Horne et al., 1982; Rosenberg et al., 1990). However, unlike these previous studies, we found no significant difference between patients diagnosed with schizophrenia and controls for the repetition factor. This has probably to do with the low frequency of repetition biases in our samples. Also contrary to our expectation, we did not find significant RNG differences between the young adult and mid-age healthy controls. Apparently, our mid-age subsample was too young and too healthy to find subtle deficits in central executive resources with the RNG parameters.

GENERAL DISCUSSION

This study replicated previous findings with the RNG task, but also added new data about the psychometric properties of the RNG task. More specifically, the present studies examined factor structure, test-retest reliability, construct validity, and criterion-related validity of the Ginsburg and Karpiuk (1994) RNG indices in samples of young adults, mid-age adults, and patients diagnosed with schizophrenia.

Our extracted factors resemble those from previous RNG factor analyses (Friedman & Miyake, 2004; Miyake et al., 2000; Towse & Neil, 1998) using Towse and Neil's RgCalc program indices. In previous studies, PCA identified three uncorrelated factors, with the first factor loading on randomness indices similar to our seriation factor (i.e., indices that are sensitive to the degree to which stereotype sequences are produced, named prepotent associates). The second factor had high loadings for indices showing clear similarities with our cycling factor (i.e., indices assessing the degree to which each number is produced at the same frequency, named "equality of response usage"). Factor three was described by Friedman and Miyake (2004) as repetition avoidance, which is similar to our repetition factor.

The test-retest correlations of the RNG scales in healthy controls and patients diagnosed with schizophrenia failed to reach the minimum of .80 required for a sound clinical tool (Anastasi & Urbina, 1997). When comparing the test-retest reliability of the RNG to more traditional, well-studied executive function tasks like the WCST (see for example Heaton et al., 1993), its stability is modest. However, no substantial practice effects were found on RNG factor scores of healthy controls and patients diagnosed with schizophrenia. Because our study was one of the first to explore test-retest stability of the RNG, future studies should further shed light on this issue, using larger samples of clinical patients and healthy participants over various periods of time (e.g., two weeks *vs*. six months).

As hypothesized, seriation was found to be related to processing speed and interference susceptibility in healthy controls and general executive functioning in the schizophrenic sample. In this clinical group, poor executive functioning was also associated with the cycling factor. Thus, it seems that RNG indices loading on the seriation and cycling factors measure deficits in executive or "frontal" functions, possibly originating from psychopathology or neurological deficits. For the repetition factor, floor effects may explain why this factor was not associated with other neurocognitive tasks. In both the healthy and the clinical sample, correlations between most RNG factors and various neurocognitive tasks were moderate. Future studies should relate Ginsburg and Karpiuk's factors to other neurocognitive tasks

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

to further establish construct validity of the RNG, or conduct latent variable analyses to see whether these factors relate to a more unitary executive function or represent independent executive subprocesses (e.g., Miyake et al., 2000).

Over the past years, several cognitive and structural theoretical models for explaining RNG deviations from randomness have been introduced, such as the aleatory model (Treisman & Faulkner, 1987), the network modulation model (Jahanshahi et al., 1998), the Wagenaar model (1970, 1972), and the Baddeley model (1986; Baddeley et al., 1998). A detailed description of these models is beyond the scope of this paper. However, what these models share is that they converge on the notion that RNG is attention demanding and reflects the limited capacity of central executive working memory and other executive functions (but see Treisman & Faulkner for a signal-detection based model), needed to suppress stereotyped sequences (inhibition) and to track and update recent responses (monitoring output) (see Baddeley, 1986; Baddeley et al., 1998; Jahanshahi et al., 1998). The neural substrate underlying RNG is most likely a network encompassing primarily the left dorsolateral prefrontal cortex (e.g., Jahanshahi et al., 1998). Thus, RNG is considered to be at the controlled end of the controlled-automatic continuum (see also Jahanshahi et al., 2006). The lack of practice-related improvement between the two RNG sessions in our second study further emphasizes the key role of controlled executive functioning (see also Jahanshahi et al., 2006). The data presented in this manuscript give some tentative evidence that at least three different subfunctions contribute to RNG and that not only externally induced response pace, but also individual differences in speed of processing affect the production of random series.

The limitations of the current studies deserve some comment. To begin with, given that our samples consisted largely of undergraduate students, most of whom were women, our samples had specific age constraints. Similarly, our studies relied on a highly specific clinical sample (i.e., patients diagnosed with schizophrenia), and so the usefulness of our data for normative purposes in clinical practice is limited. The effect of medication on randomization in clinical samples would also deserve further attention. Another limitation of psychometric studies like the present one is the multiple statistical testing, which raises the probability of experimenter-wise errors. Where possible and appropriate, we tried to reduce that probability by applying Bonferroni corrections. Also our studies can best be seen as a first step and the next steps could involve experimental manipulation (e.g., by dual tasks) of the RNG factors and their correlates that we identified. In future research, it may also be worthwhile to determine discriminant validity pertaining to constructs such as global intelligence and simple sustained attention. A final limitation of our studies is that we employed the 1 sec condition of the RNG task, which differs from the 1.5 sec condition in the Ginsburg and Karpiuk study (1994); but see Jahanshahi et al., (2006). Indeed, parametric research in which response pace times are systematically varied in different samples (i.e., healthy and clinical) might be informative.

Summing up, the RNG task appears to be a promising task to measure inhibition, updating, and monitoring functions in normal as well as clinical populations. Failures in these functions are reliably tapped by the RNG task. Although it does not (yet) possess the psychometric properties of a clinical tool, as a research tool the RNG may help us understand nonrandom response biases in healthy humans and even more prominent deviations from randomness in clinical populations.

ACKNOWLEDGMENTS

This study was supported by grant number 452-02-006 from the Netherlands Organization for Scientific Research (N.W.O.). The authors thank Nicole Haas, Hilde Verbeek, Renske Rigter, and Marije de Vos for help in collecting the data. The information described in this manuscript and the manuscript itself is new and original and has never been published either electronically or in print, and there is no conflict of interest, either financial or other.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association Press.

Anastasi, A. & Urbina, S. (1997). *Psychological testing* (7th ed). Englewood Cliffs, NJ: Prentice Hall.

Artiges, E., Salamé, P., Recasens, Ch., Poline, J.B., Attar-Levy, D., De la Raillère, A., Paillère-Martinot, M.L., Danion, J.M., & Martinot, J.L. (2000). Working memory control in patients with schizophrenia: A PET study during a random number generation task. *American Journal of Psychiatry*, *157*, 1517–1519.

Axmacher, I., Bente, D., & Ferner, U. (1970). Informationsstatistische Untersuchungen zur Struktur einfacher Handlungsfolgen bei endogenen Psychosen (Statistical examination of the structure of simple action sequences in endogene psychosis). *Arzneimittel-Forschung* (*Drugs Treatment Research*), *20*, 919–921.

Baddeley, A.D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, *18*, 119–129.

Baddeley, A.D. (1986). *Working memory*. Oxford, England: Oxford University Press.

Baddeley, A.D. (1998). The central executive: A concept and some misconceptions. *Journal of the International Neuropsychological Society*, *4*, 523–526.

Baddeley, A.D., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and executive control of working memory. *Quarterly Journal of Experimental Psychology*, *51A*, 819–852.

Bradshaw, J.L. & Mattingley, J.B. (1995). *Clinical Neuropsychology: Behavioral and brain science*. San Diego: Academic Press.

Brown, R.G., Soliveri, P., & Jahanshahi, M. (1998). Executive processes in Parkinson's disease: Random number generation and response suppression. *Neuropsychologia*, *36*, 1355–1362.

Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual and Motor Skills*, *84*, 627–661.

Brugger, P., Monsch, A.U., Salmon, D.P., & Butters, N. (1996). Random number generation in dementia of the Alzheimer type: A test of frontal executive functions. *Neuropsychologia*, *34*, 97–103.

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press

Brugger, P., Pietzsch, S., Weidmann, G., Biro, P., & Alon, E. (1995). Stroop-type interference in random number generation. *Psychological Reports*, *77*, 387–390.

Daniels, C., Witt, K., Wolff, S., Jansen, O., & Deuschl, G. (2003). Rate dependency of the human cortical network subserving executive functions during generation of random number series: A functional magnetic resonance imaging study. *Neuroscience Letters*, *345*, 25–28.

De Zubicaray, G.I., Smith, G.A., Chalk, J.B., & Semple, J. (1998). The modified card sorting test: Test-retest stability and relationships with demographic variables in a healthy older adult sample. *British Journal of Clinical Psychology*, *37*, 457–466.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage Publications.

Friedman, N.P. & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent–variable analysis. *Journal of Experimental Psychology: General*, *133*, 101–135.

Gerton, B.K., Brown, T.T., Meyer-Lindenberg, A., Kohn, P., Holt, J.L., Olsen, R.K., & Berman, K.F. (2004). Shared and distinct neurophysiological components of the digits forward and backward tasks as revealed by functional neuroimaging. *Neuropsychologia*, *42*, 1781–1787.

Giesbrecht, T., Merckelbach, H., Geraerts, E., & Smeets, E. (2004). Disruptions in executive functioning and dissociation in undergraduate students. *Journal of Nervous and Mental Disease*, *192*, 567–569.

Ginsburg, N. & Karpiuk, P. (1994). Random number generation: Analysis of responses. *Perceptual and Motor Skills*, *79*, 1059–1067.

Ginsburg, N. & Karpiuk, P. (1995). Simulation of human performance on a random generation task. *Perceptual and Motor Skills*, *81*, 1183–1186.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual: Revised*. Odessa, Florida: Psychological Assessment Resources.

Horne, R.L., Evans, F.J., & Orne, M.T. (1982). Random number generation, psychopathology and therapeutic change. *Archives of General Psychiatry*, *39*, 680–683.

Jahanshahi, M., Profice, P., Brown, R.G., Ridding, M.C., Dirnberger, G., & Rothwell, J.C. (1998). The effects of transcranial magnetic stimulation over the dorsolateral prefrontal cortex on suppression of habitual counting during random number generation. *Brain*, *121*, 1533–1544.

Jahanshahi, M., Saleem, T., Ho, A.K., Dirnberger, G., & Fuller, R. (2006). Random number generation as an index of controlled processing. *Neuropsychology*, *20*, 391–399.

Jelicic, M., Henquet, C.E.C., Derix, M.M.A., & Jolles, J. (2001). Test-retest stability of the behavioural assessment of the dysexecutive syndrome in a sample of psychiatric patients. *International Journal of Neuroscience*, *110*, 73–78.

Joppich, G., Däuper, J., Dengler, R., Johannes, S., Rodriguez-Fornells, A., & Münte, T.F. (2004). Brain potentials index executive functions during random number generation. *Neuroscience Research*, *49*, 157–164.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

Krabbendam, L. & Kalff, A.C. (1998). *The behavioural assessment of the dysexecutive syndrome–Dutch version*. Lisse, The Netherlands: Swets & Zeitlinger.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Peters, M.J.V., Jelicic, M., Haas, N., & Merckelbach, H. (2006). Mild executive dysfunctions in undergraduates are related to recollecting words never presented. *International Journal of Neuroscience*, *116*, 1065–1077.

Rabinowitz, F.M. (1970). Characteristic sequential dependencies in multiple-choice situations. *Psychological Bulletin*, *74*, 141–148.

Rosenberg, S., Weber, N., Crocq, M.A., Duval, F., & Macher, J.P. (1990). Random number generation by normal, alcoholic and schizophrenic subjects. *Psychological Medicine*, *20*, 953–960.

Salamé, P., Danion, J.M., Peretti, S., & Cuervo, C. (1998). The state of functioning of working memory in schizophrenia. *Schizophrenia Research*, *30*, 11–29.

Schneider, S., Joppich, G., van der Lugt, A., Däuper, J., & Münte, T.F. (2004). Brain potentials and self-paced random number generation in humans. *Neuroscience Letters*, *367*, 51–55.

Shinba, T., Shinozaki, T., Kariya, N., & Ebata, K. (2000). Random number generation deficit in schizophrenia characterized by oral vs. written response modes. *Perceptual and Motor Skills*, *91*, 1091–1105.

Stevens, J.P. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Stinissen, J., Willems, P., Coetsier, P., & Hulsman, W. (1970). *Handleiding bij de Nederlandstalige bewerking van de Wechsler Adult Intelligence Scale (WAIS)*. [*Manual to the Dutch version of the Wechsler Adult Intelligence Scale (WAIS)*]. Lisse, The Netherlands: Swets & Zeitlinger.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Towse, J.N. (1998). On random generation and the central executive of working memory. *British Journal of Psychology*, *89*, 77–101.

Towse, J.N. & Neil, D. (1998). Analyzing human random generation behaviour: A critical review of methods used and a computer program for describing performance. *Behavior Research Methods, Instruments & Computers*, *30*, 583–591.

Treisman, M. & Faulkner, A. (1987). Generation of random sequences by human subjects: Cognitive operations or psychophysical process? *Journal of Experimental Psychology: General*, *116*, 337–355.

Van der Linden, M., Beerten, A., & Pesenti, M. (1998). Age-related differences in random number generation. *Brain and Cognition*, *38*, 1–16.

Verhage, F. (1964). *Intelligentie en leeftijd* (*Intelligence and age*). Assen, The Netherlands: Van Gorcum.

Wagenaar, W.A. (1970). Subjective randomness and the capacity to generate information. *Acta Psychologica*, *33*, 233–242.

Wagenaar, W.A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, *77*, 65–72.

Williams, M.A., Moss, S.A., Bradshaw, J.L., & Rinehart, N.J. (2002). Random number generation in Autism. *Journal of Autism and Developmental Disorders*, *32*, 43–47.

Wilson, B.A., Evans, J.J., Emslie, H., Alderman, N., & Burgess, P. (1998). The development of an ecologically valid test for assessing patients with a dysexecutive syndrome. *Neuropsychological Rehabilitation*, *8*, 213–228.

https://doi.org/10.1017/S1355617707070786 Published online by Cambridge University Press