British Journal of Developmental Psychology (1999), 17, 363–380 *Printed in Great Britain*

An exploration of random generation among children

John N. Towse*
Royal Holloway University of London, UK

Amy Mclachlan
Brunel University, UK

The generation of random sequences is known to be a complex, demanding and effortful task for adults. This study explores random generation performance among children in three experiments. Expt 1 illustrates 8–10-year-olds' sensitivity to response speed requirements. Expt 2 shows that 8–11-year-olds were sensitive to the number of response alternatives, while there was equivalence in output quality over two types of instructional formats. Expt 3 reveals competencies in performance among 5–7-year-olds and shows that response repetitions are partly amenable to instructional emphasis. Across comparable studies, analysis confirmed a multidimensional structure to response sets. Generally, data show the potential utility of random generation as a developmental task with substantial and multifaceted attentional requirements.

The term 'working memory' has come to be used as both a conceptual heuristic (emphasising the importance of memory phenomena for cognitive actions) and a theoretical model. According to Baddeley's (1986; 1996) specific model, working memory comprises multiple components with two memory slave systems, a 'phonological loop' and a 'visuospatial sketch pad', governed by a central coordinating structure, a 'central executive'. Despite its austerity of specification, the working memory model, with its divide-and-rule strategy, has been remarkably successful in accounting for a disparate range of cognitive phenomena. It has also provided an interesting and useful interpretative framework for developmental research (Halliday & Hitch, 1988).

Through the concept of a central executive, this model of working memory makes task control an important issue. However, the structural framework is not very forthcoming with regards to *how* such control is exercised, or the conditions under which it is successful or unsuccessful (though see Norman & Shallice, 1986). The issue is particularly pertinent for developmental research, since children may have the same basic mental structures and systems as adults, while lacking the strategic and integrative abilities that permit the flexibility, or diversity of representation, that are the hallmarks of adult cognition. This

* Requests for reprints should be addressed to John N. Towse, Department of Psychology, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK.

possibility is given impetus by the increasing interest shown in the notion that developmental change is associated with variation in inhibitory control (Dempster, 1993; Diamond, Prevor, Callender & Druin, 1997). That is, one marker of effective cognition is the ability to respond to current circumstances, and not just to be tethered to those actions and responses that have been successful on previous occasions or trials. Inhibition of responses that are prepotent by virtue of immediately preceding eliciting conditions (a chain of actions), or those that are generally of a low threshold, becomes an important cognitive operation.

One cognitive task that appears to require inhibitory control to a large extent is random sequence generation. Commonly, participants must produce a long sequence of items from a restricted vocabulary so that responses form a random, or unordered, set. For successful task performance, individuals must overcome (inhibit) a tendency to produce 'natural' (i.e. associated or chain-like) patterns. On the basis of adult data, Baddeley (1986) has suggested that randomization might be a prototypical 'executive' task fully stretching the central executive system. In this respect, this study considers the characteristics of a putative executive task in children.

It is clear that random generation is highly demanding for adults, with substantial deviations from randomness in most data sets (Wagenaar, 1972; but see Neuringer, 1986). Individual–difference analysis (Ginsburg & Karpiuk, 1994; Towse & Neil, 1998) and experimental analysis suggest that multiple factors underlie performance. Concurrent task performance (Salway, 1991; Spence, Wolitzky & Pezenik, 1969), increased time stress (Baddeley, 1966; Robertson, Hazlewood & Rawson, 1996) and a larger response set (Warren & Morin, 1965; Wiegersma, 1976) all impair performance. The avoidance of immediate item repetition is also common (Brugger, 1997). So, while imperfect concepts of randomness also constrain performance (Bar-Hillel & Wagenaar, 1991), producing random sequences is evidently complex. Thus, developmental change might occur along several different dimensions, while other aspects of performance might be age-invariant.

An early and rare study of randomization among children was reported by Thomas (1969), using a specific and rather idiosyncratic index of performance. Response production rate and the ability to choose alternatives with equal frequency increased between approximately 6 and 18 years of age, though most children studied were between 10 and 17 years old. Rabinowitz, Dunlap, Grant & Campione (1989) reported developmental improvement in random generation, too. However, Rabinowitz *et al.* (1989) appeared to be primarily interested in producing a mathematical, rather than psychological, model of random generation. In addition, some performance manipulations were rather unorthodox; in one condition, there was a 10 s interval between each response. The relevance of task length, boredom and concentration over large unfilled intervals (particularly, perhaps, for younger children) appear non-trivial methodological concerns.

Given the relative scarcity of data on children's randomization abilities, this study sought to characterize their performance in a number of ways. There follows a series of experiments that focused on children between 4 and 11 years of age. This represents an important period of strategic development (Bjorklund, 1990) and there is evidence for developmental changes in 'executive'-type tasks in this age band, accompanying the neurological maturation of frontal lobe functioning (Levin *et al.*, 1991; Welsh, Pennington & Groisser, 1991). The present studies evaluate the feasibility of administering randomization instructions to children. Performance is measured using several different indices, giving greater breadth to response assessment, and providing a bridge to adult studies incorporating these indices. Finally, experimental manipulations of response speed, response vocabulary size and instructional content are examined.

EXPERIMENT 1

This first study examines children's ability to produce random sequences under different response speeds. Conventionally, studies of random generation elicit (at least) 100 responses. To limit the impact of task length and boredom among children, 70 responses were collected here.

Method

Participants and design

Analysis was based on 42 children from two school classes. There were 25 younger children, mean age 8 years 8 months (ranging 8;3 to 9;2) and 17 older children, mean age 10 years 9 months (ranging 10;3 to 11;2). Age formed the single between-participants factor, while response speed ('slow' or 'fast') was a within-participants manipulation. Half the children performed a separate task (not described here) before generating number sequences. Data from one child with an independent assessment of autistic-spectrum behaviour were excluded.

Procedure

Children were tested individually in a quiet area of school, with the task introduced as a game involving numbers. The experimenter asked children to generate (verbally) a list of numbers between 1 and 10. They were asked: '. . . to jumble up the numbers as much as possible. So the way I'd like you to say the numbers is as if you were rolling a dice in your head [at this point, a six-sided die was produced and rolled several times]. So you might roll a "6", and then roll a "2", and so on. That is how I would like you to say the numbers. So, if you are going to produce a number like a dice, you won't want to say the same number all the time, or say numbers that follow after each other, like "2", then "3", then "4". A dice isn't likely to give patterns like that. Every time you hear a beep just say a number between 1 and 10 trying to make them jumbled up just like the dice would.'

A production rate of 1.5 s per response ('fast') and 2.5 s per response ('slow') was employed in counterbalanced order. Times, slightly slower than common in adult studies, reflected a pace that was demanding yet practicable. In each set, 70 responses were elicited.

Results

Occasional response failures (e.g. an absence of response in the allotted time) occurred, but these were infrequent and were not recorded systematically. Response regularities can be quantified in (an indefinite number of) different ways. Indeed, because sequences can be biased by all manner of production methods (e.g. by using only a subset of responses, or repeatedly producing particular combinations), it is important to have different measures of performance. The following analytic approach was, therefore, adopted. Adult data have indicated four substantial factors influencing random sequences (Towse & Neil, 1998), a finding corroborated by children's responses to be reported later. So four scores that loaded highly on these factors were entered as dependent variables into a doubly-multivariate analysis of variance. Univariate follow-up tests used Bonferroni correction, making analysis of specific measures quite conservative. Measures of effect size are reported throughout.

Table 1. Mean randomness scores in Expt 1 (SD in parentheses)

Age group	8 years		10 years	
Response speed	slow	fast	slow	fast
Randomness measure:				
Distributions (R)	3.15 (1.86)	2.29 (1.60)	2.42 (1.37)	4.18 (5.36)
Chain lengths (TPI)	72.2 (15.0)	67.1 (13.7)	84.7 (11.4)	76.3 (14.5)
Immediate repetitions (Phi2)	−3.75 (0.56)	−3.33 (1.15)	−3.89 (0.27)	−4.14 (0.72)
Distant repetitions (Phi7)	−1.64 (1.09)	−1.96 (1.07)	−0.92 (2.21)	−1.37 (2.08)

The dependent variables were R, TPI, Phi2 and Phi7. These variables assess, respectively, the evenness of response alternative usage (the sampling distribution), the extent to which ascending and descending runs or chains are used, immediate repetitions and delayed repetitions of responses (see Appendix 1 for further descriptions)[1]. For convenience the variables are labelled 'distributions', 'chain lengths', 'immediate repetitions' and 'distant repetitions'. Mean scores are detailed in Table 1. Analysis with age and response speed as factors, indicated a significant age difference in overall performance ($F(4,37) = 2.92, p < .05$, multivariate $\eta^2 = .240$), and greater randomness at the slower response speed ($F(4,37) = 2.74, p < .05$, multivariate $\eta^2 = .229$). There was a near-significant interaction between these factors, ($F(4,37) = 2.55, p < .10$, multivariate $\eta^2 = .216$).

Univariate tests indicated that the main effect of age group occurred largely because older children used smaller sequence chain lengths than younger children ($F(1,40) = 9.35, p < .05$, partial $\eta^2 = .18$). To a lesser extent, age groups differed also by immediate repetitions ($F(1,40) = 5.71, p < .10$, partial $\eta^2 = .125$), with older children showing more extreme repetition avoidance. The main effect of response speed was largely attributable to the measure of chain lengths, with longer ordinal sequences at the fast speed ($F(1,40) = 7.04, p < .05$, partial $\eta^2 = .15$). The interaction trend between age and speed centred on response distributions and immediate repetition scores ($F(1,40) = 5.37$, partial $\eta^2 = .118$; $F(1,40) = 5.64$, partial $\eta^2 = .124$ respectively), though in neither case was the effect significant after Bonferroni adjustment. There was little evidence for an interaction between age and speed for chain lengths ($F < 1$, partial $\eta^2 = .011$).

The recourse to stereotyped 'count' responses is illustrated in Fig. 1, showing the frequencies of arithmetic differences between each response and its preceding neighbour. The strong avoidance of immediate repetitions is evident, and is illustrated further in Fig. 2, a frequency distribution of the lag, or gap, between item repeats (providing a simpler, graphic variant of Phi scores, but over a greater repetition range). Once emitted, responses are initially avoided. This 'inhibition' dissipates gradually over intervening choices and

[1] Randomization research commonly makes reference to RNG scores, a measure of redundancy in paired sequences. RNG scores showed effects in the same direction as TPI. Since TPI scores load more highly on the major factor in principal components analysis, it was used as a dependent variable here.

not merely as a function of elapsed time, since this would produce a marked leftward shift in the fast response distribution. The modal repetition distance lag is around six steps, with similar profiles for both response speeds and age groups. The difference between 8- and 10-year-olds at a lag of six items for the fast speed was not significant ($t(26.5) = .85$, $\eta^2 = .017$).

Examining the relationships between measures, there were negative correlations between response distributions and immediate repetitions, and chain lengths and immediate repetitions ($r(40) = -.40$ and $-.50$, $ps < .01$ (even distributions and long sequence lengths associated with less repetition avoidance)). There was also a positive correlation between distributions and distant repetitions ($r(40) = .47, p < .01$), such that a more even response distribution was associated with fewer repetitions with a lag of six items.

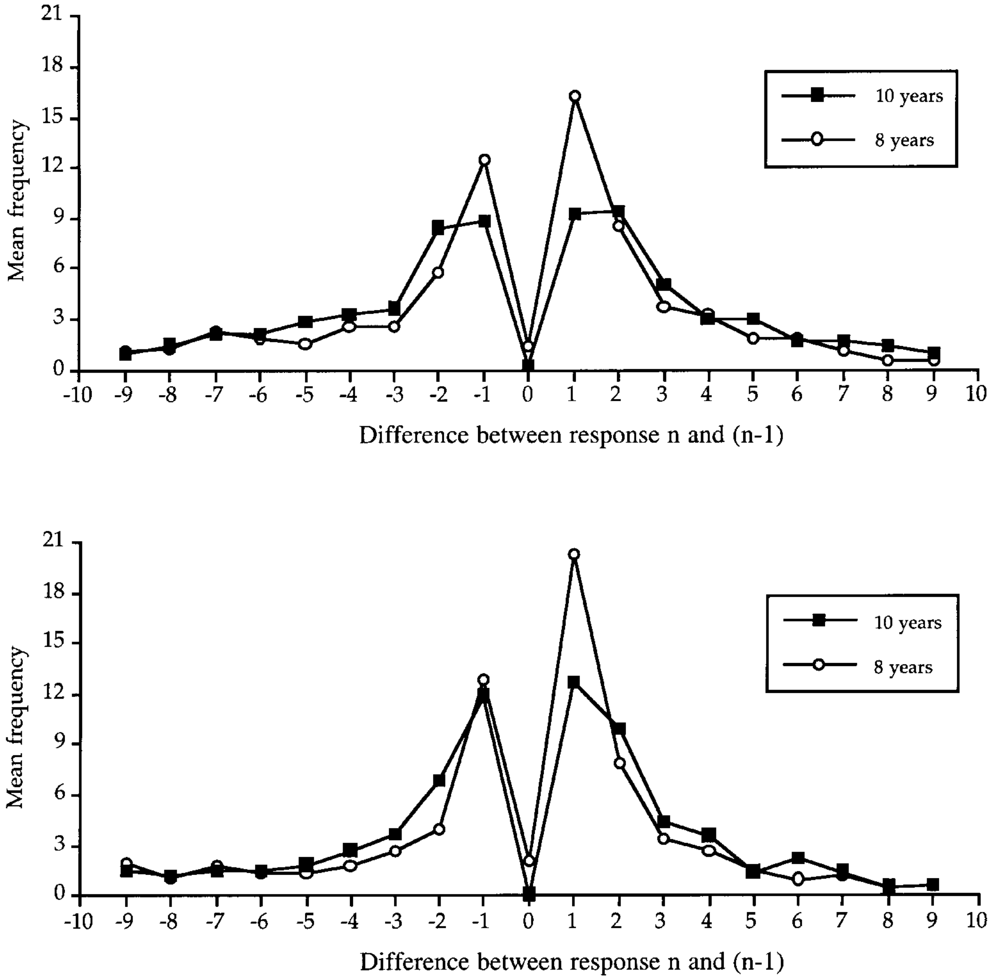

Figure 1. Distribution of first order difference scores in Expt 1 – the numerical difference between successive responses. Upper panel represents responses in the slow condition, lower panel represents responses in the fast condition.

Discussion

Data from this experiment lay the foundations for an understanding of children's random generation in a number of ways. Children aged 8 and 10 years are, in the main, quite able to attempt random number production, while they fall short of ideal random generators in several respects. Nonetheless, their responses indicate that they approach the task in the same general fashion, and that children fare less well with a reduced time interval in which to prepare a response.

Although there is a developmental improvement in performance, most prominently with respect to chain lengths (increasing and decreasing numeric series), the size of the response speed effect with respect to chain lengths is rather invariant across age. Clearly,

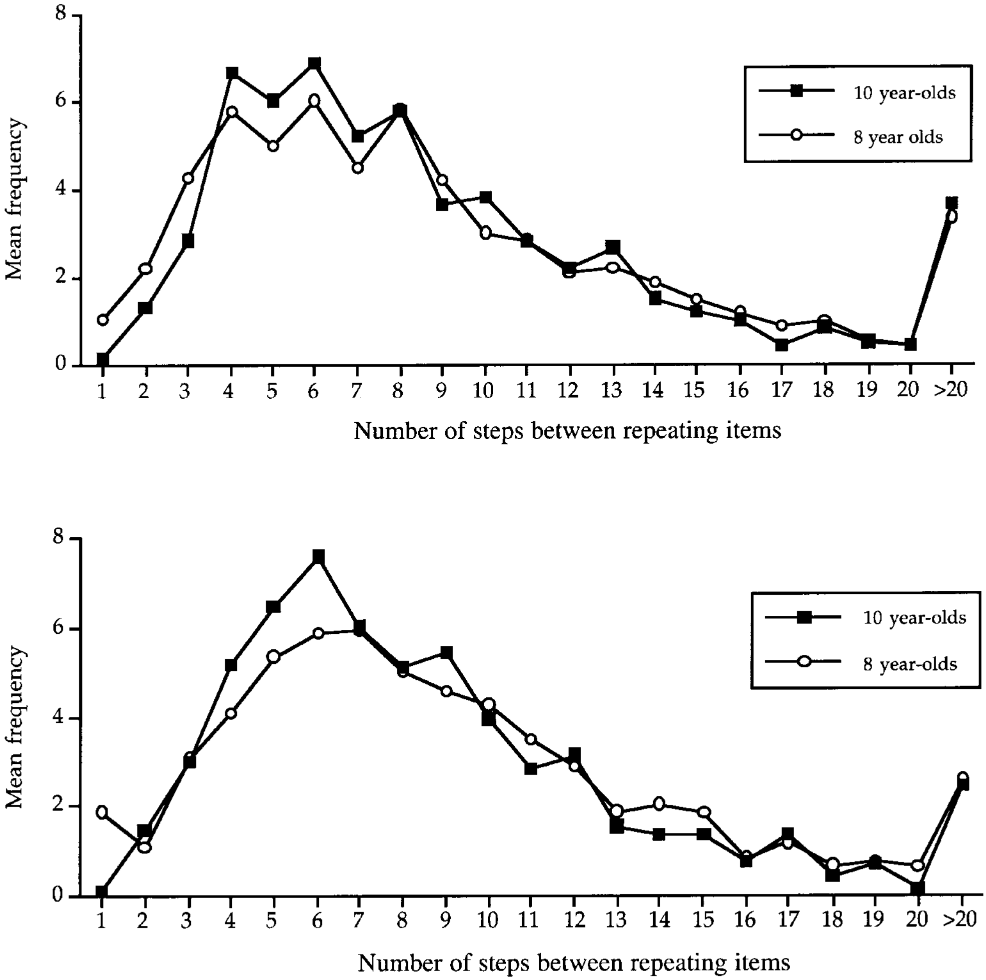

Figure 2. Distribution of distances between repeating items in Expt 1. For all response alternatives, the difference in the response positions for each repetition is determined and summed. Upper panel represents responses in the slow condition, lower panel represents responses in the fast condition.

the lack of an age by speed interaction should be interpreted with considerable caution, being a null effect. One possibility, though, is that the age difference represents more than a simple change in overall 'processing capacity'. If that were the case, manipulating task demand should produce a smaller proportional effect among older children (those with a larger capacity).

For response distributions, there were trends for older children to show a decline in randomness at the faster response speed, while younger children became more random. Among adults, R scores are sometimes insensitive to speed manipulations (see e.g. Towse, 1998). It seems appropriate, therefore, to conclude that time stress primarily impairs the avoidance or inhibition of prepotent associations, leading to adjacent sequence pairs (see Fig. 1) or long ordinal runs of responses (chain lengths). Distributing responses appropriately over the range of legitimate values (the aspect of randomization measured by R scores) takes place largely independently of this inhibition effect. The speed-related improvement in R quality among younger children might then be seen as a by-product of the way associated sequences are emitted. That is, long strings of adjacent values (1, 2, 3, etc.) will distribute responses evenly across alternatives.

EXPERIMENT 2

Age differences in Expt 1 could be explained by a developmental increase in the efficiency of the cognitive system(s) subserving random generation. Alternatively, younger children may interpret the instructions less appropriately. That is, age differences might be an artifact of how children construe task demands. While it would be difficult to examine all possible task instructions, it would be reassuring to confirm that particular experimental frames are not crucial for responses. Thus, alternative instructions were prepared, involving a 'hide-and-seek' type game. Since this situation is likely to be familiar (perhaps more familiar than the details of rolling dice), the variation in instructions provide an opportunity to evaluate whether the wording used to present the task is critical.

Method

Participants and design

Data were available for 36 children from two school classes. There were 18 younger children, mean age 8 years 5 months (ranging 8;4 to 8;6), and 18 older children, mean age 11 years 2 months (ranging 11;0 to 11;3). Age and experimental scenario ('dice' or 'rooms' instructions) were between-participants factors, while response set size ('small' or 'large') was a within-participants manipulation.

Procedure

Children were tested individually in a quiet area of school, producing 70 responses (one response every 2.5 seconds), using the range 1–7 and 1–10 inclusive, in counterbalanced order. The verbal instructions were as follows:

> *Dice instructions*: 'Let's play a game with numbers. What I want you to do is pretend that you are rolling a dice, like this one here, and then tell me what number appears when you've rolled it. The dice that you can roll has all the numbers between 1 and (7)(10). So each time you hear a tone from the tape recorder, say the number on the dice that you have pretended to roll.
>
> So what I want you to do is try and say numbers just like a dice would produce numbers, without any order to them. So if you said number 2, then number 3, then number 4, then number 5, then that

wouldn't be much like the roll of a dice would it? So what you should try and do is jumble up the order.

You can use all the numbers between 1 and (7)(10), and you can choose the same number twice by just repeating it if you want to. But remember to tell me which number you have chosen each time you hear the signal.

So you might say number 4, then number 1, then number 6, then number 3, then number 3 again, and so on.'

Rooms instructions: 'Let's play a pretend game of hiding. What I want you to do is pretend you are hiding from someone who is out to catch you. You have (7)(10) rooms that you can hide in, and you have to keep telling me which room you are going to pick, by saying its number each time you hear a tone from the tape recorder.

Now, if you are going to try and hide, you should choose rooms where the person won't look. The person trying to find you will know where you have been, so you will just have to keep fooling him!

So if you choose room 2, then room 3, then room 4, then room 5, the person chasing you is likely to guess where you will go next – to room 6. So what you should try and do is choose a room number that no-one else would expect.

You can use all the rooms between 1 and (7)(10), and it takes the same time to go from any room to any other. You can also choose to stay in the same room by saying the same number again. But remember to tell me which room you are going to each time you hear the signal.

So you might choose room 4, then room 1, then room 6, then room 3, then room 3 again, and so on.'

Results

In this experiment, before the experimental conditions could be legitimately compared, randomization scores require adjustment. As an illustration of why this is necessary, note that with more response alternatives, adjacent pairs should become relatively less frequent in random sequences, as there are alternative permutations available. To control for the natural variation in 'baseline' randomness values, Monte Carlo computer simulations were run on 2000 quasi-random sets of 70 responses to produce theoretic mean and standard deviation values. Children's sequences were then described by a z-score statistic for each dependent variable, representing the degree to which the sequence deviated from a random sequence. Table 2 presents the normalized scores.

Multivariate analysis of variance showed no main effect of age ($F(4,29) = 1.76$, multivariate $\eta^2 = .195$), nor a reliable effect of instruction ($F(4,29) = 2.62$, $p < .10$, multivariate $\eta^2 = .265$), while there was a strong effect of response set size, ($F(4,29) = 18.3$, $p < .01$, multivariate $\eta^2 = .720$). None of the interactions was significant (age $\times$ instruction, $F(4,29) = 1.81$, multivariate $\eta^2 = .200$; age $\times$ set size, $F(4,29) = 1.17$, multivariate $\eta^2 = .139$; instruction $\times$ set size, $F < 1$, multivariate $\eta^2 = .068$; age $\times$ instruction $\times$ set size, $F < 1$, multivariate $\eta^2 = .094$).

Univariate tests confirmed that responses were less random for the larger set size in terms of chain lengths (ascending/descending sequences; $F(1,32) = 60.8$, partial $\eta^2 = .655$), while the avoidance of immediate repetitions was less severe with a larger response set ($F(1,32) = 18.6$, partial $\eta^2 = .369$. In fact, absolute repetition frequency was comparable in the two conditions, $t(35) = .61$, $p > .05$, $\eta^2 = .011$. All remaining effects were non-significant. However, prior to Bonferroni correction, the rooms instructions produced more immediate repetitions ($F(1,32) = 4.83$, partial $\eta^2 = .131$), and the older children tended to produce a more even response distribution ($F(1,32) = 4.66$, partial $\eta^2 = .127$).

Table 2. Mean randomness scores as *z*-score deviations in Expt 2 (SD in parentheses)

Age group	8 years		11 years	
Number of responses	7	10	7	10
Randomness measures with 'dice' instructions:				
Distributions (R)	0.64 (1.62)	1.99 (1.89)	−0.78 (0.40)	−0.66 (0.69)
Chain lengths (TPI)	−0.10 (1.15)	−1.51 (1.49)	0.76 (1.20)	−1.02 (1.48)
Immediate repetitions (Phi2)	−3.21 (0.30)	−2.70 (0.56)	−3.16 (0.22)	−2.45 (0.28)
Distant repetitions (Phi7)	−0.72 (1.14)	−1.04 (1.61)	−0.91 (0.92)	−0.30 (1.22)
Randomness measures with 'rooms' instructions:				
Distributions (R)	−0.20 (1.47)	0.05 (0.83)	−0.20 (1.97)	−0.22 (2.03)
Chain lengths (TPI)	−0.16 (2.04)	−2.30 (2.07)	0.28 (1.23)	−1.34 (1.64)
Immediate repetitions (Phi2)	−1.98 (2.16)	−1.83 (1.55)	−2.95 (0.47)	−2.18 (0.75)
Distant repetitions (Phi7)	−0.46 (0.73)	−0.29 (0.58)	−0.37 (1.19)	−0.07 (1.57)

Discussion

This experiment confirms that, as for adults (Towse, 1998; Warren & Morin, 1965), children's responses are sensitive to the number of response alternatives. Data also show that repeats are avoided even when instructions reinforce their legitimacy (although there was a non-significant trend for the rooms scenario to elicit more repetitions).

In general, the similarity of random generation performance across instructions lends support to the view that responses reflect children's attempts at avoiding pattern or order (i.e. measures their attempts at being random). That is not to say that instructions are unimportant for performance generally, or for particular aspects of response style. Nonetheless, in testing children on randomization, providing a scenario that makes sense to the child is probably more important than the particular story concerned.

EXPERIMENT 3

While the preceding experiments portray children's randomization in a number of ways, an important question that remains is the lower age boundary at which the task can be usefully administered to children. That is, one may ask: Is there a period below which children do not produce meaningful responses, and is this boundary imposed by constraints of performance (the task becomes too difficult) or competence (the task is not understood)?

On close analysis, this question is not easy to answer precisely, especially given that the performance/competence distinction is a notoriously elusive one (for one recent perspective, see Sophian (1997) and following commentaries). Randomization is perhaps particularly illustrative of this problem because in producing responses, there is no formally correct or incorrect answer. Although *sequences* can be graded for order, individual *items* are ambiguous. In addition, since sequences are evidently performance-sensitive, no sharp developmental demarcation line can be universally expected to separate acceptable from unacceptable response sets. Acknowledging these caveats, the randomization task was presented to a group of children between 4 and 7 years of age to evaluate whether the task can be completed by children younger than those used hitherto. Responses were required at a relatively fast rate so as to explore whether children can conform to the required rhythmic demands and to yield a conservative estimate of ability. The study also explored the possibility that instructions emphasizing the appropriacy of repetitions would alter responses, thus examining the inevitability of repetition avoidance behaviour shown in preceding studies.

Method

Participants and design

Thirty-eight children from three classes at an infant school were invited to play a game of thinking up numbers. Analysable data were available for 31 children, ranging in age from 4;11 to 7;6. As part of the task instructions, the emphasis placed on the production of repetitions formed a between-participants manipulation with two levels, a 'standard' and a 'repetition emphasis' condition.

Procedure

Children were tested individually in a quiet area of school and produced 70 responses at the rate of one number every 1.5 s, using the numbers 1–10 (inclusive). The general format for explaining the task corresponded to the dice instructions on previous studies. The experimenter followed a general script by showing a 10-sided die and asking: 'Do you know what this is? That's right, it's a dice, though it is a bit strange, because it has more faces and more numbers than a dice normally does. Anyway, let's roll this dice and see what numbers we get.'

The experimenter then rolled the die several times and the child announced the result. The experimenter continued:

> 'Do we really know what number will come up if we roll the dice again? No, we don't, because it could be any number. Well, what I want you to do is pretend to be a dice, like this one here, and then by pretending to be a dice you can say all the different numbers that come up when the dice is rolled.
>
> When you pretend, use all the numbers between 1 and 10, and each time you hear a beep from the tape, say the number you thought of. So what I want you to do is try and say numbers just like a dice would, so that they are all jumbled up.
>
> So if you said number 1, then next you said number 2, then number 3, 4, 5, 6, 7, 8, 9, 10, would that be the way a dice would give numbers? Uh-Uh, it isn't really, because the numbers would be all jumbled up from a dice. So what you should try and do is mix up all the numbers as much as you can.'

In the standard instructions, children were reminded to say a number for each signal from the tape, and were told that they could repeat any number if they wanted to. In the repetition condition, this was emphasized and elaborated, by saying: 'Sometimes, a dice comes up with the same number again, doesn't it? You roll a "2" and the next time you roll a "2" again. So when you pretend to roll a dice in your head, sometimes, are you going to choose the same number again? Yeah? That's OK because that is what a dice would do sometimes. There might be a 4 and then another 4. So will you say the same number sometimes?'

Results

Four children (aged between 4;8 and 5;3) who were invited to produce random responses declined to produce sufficient responses (one case), produced only ordered sequences (one case), or failed to follow paced instructions to such an extent that data collection was terminated (two cases). Two older children (6;7 and 6;6) were also dropped from analysis because of excessive failure in keeping to paced instructions, while one additional child (aged 7;7) independently diagnosed with Asperger's syndrome was not included.[2]

Table 3. Correlations between age, randomness and number of task violations

	Age	TPI	RNG	R	Phi2	Phi7
TPI	.20					
RNG	−.40*	−.44*				
R	.19	.46**	.06			
Phi2	−.05	−.52**	.20	−.23		
Phi7	.06	.34^{+}	−.02	.22	.06	
Violations	−.42*	−.11	−.01	−.13	−.32^{+}	.03

$^{+}$ $.05 < p < .10$; $*p < .05$; $**p < .01$

TPI: chain lengths; RNG: Evans' randomization index; R: distributions; Phi2: immediate repetitions; Phi7: distant repetitions.

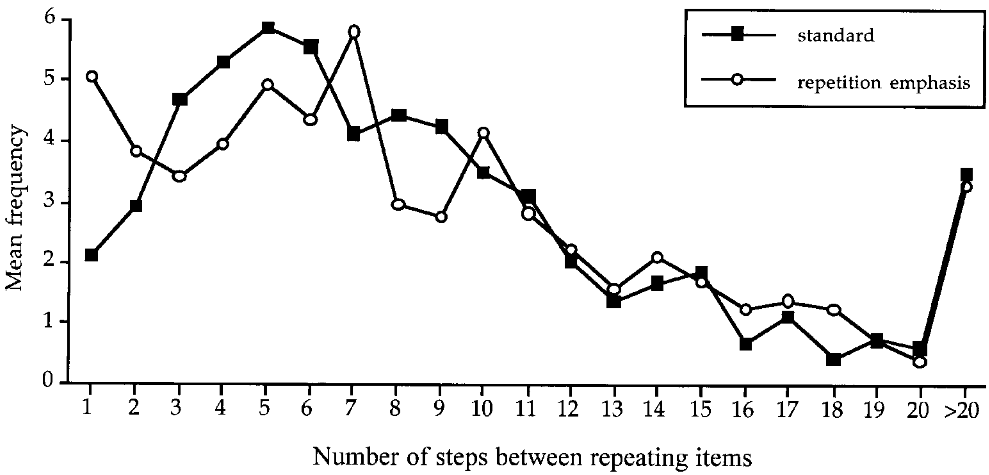

Figure 3. Distribution of repetition distances in Expt 3, following standard instructions that mention the possibility of repeating items, and repetition emphasis instructions that focus on the appropriacy of response repetitions.

A comparison of children given 'standard' or 'repetition emphasis' instructions indicated a significantly higher frequency of immediate repetitions following instructional cues ($t(22.01) = 2.27$, $p < .05$, $\eta^2 = .151$), as shown in Fig. 3. The differences between instructions at longer repetition lags were not significant. Since the study attempted to examine age changes and age limits in performance, and given the continuous distribution of age, correlational analyses were performed.

Table 3 indicates that the number of response failures (response vocabulary violations and/or missing the response cue) varied with age. Often, 5–6-year-olds produced at least

[2] Of those responses that were produced, most formed recitations of the number sequence, in tantalizing contrast to the child with autistic behaviour in Expt 1, who repeated responses on 29% of trials.

20 response failures in the set of 70 items, although only a rough estimate of task compliance was possible. Children's verbal response sometimes overlapped with the next auditory cue, making it difficult to know whether they had simply failed to respond, or were unaware that the cue had occurred (the experimenter was conservative in marking failures in this respect). Among the youngest children, whose overt articulations were slower, this problem was particularly evident, and they also tended to require more explicit prompting to continue. The correlation between response failures and age also clouds interpretation of other age analyses, since it effectively means that younger children gave slower responses, and speed is likely to affect response quality. Although the length of sequence chains did not vary significantly with age, a more focused measure of paired responses, RNG, did correlate with age ($r(29) = -.40$, $p < .05$), with young children more likely to repeat certain paired responses.

Regardless of actual performance, children were congratulated for 'pretending to be a dice' and for producing such a long string of numbers. When asked whether the task was 'hard' or 'easy', most children suggested the latter, though occasionally they echoed one comment that 'there wasn't a lot of time to say the numbers'. When asked *how* they had accomplished the task, they invariably failed to describe a process, or mentioned something to the effect that they 'said numbers that came into my head'. In summary, there appeared little insight into or identification of strategic method in performance.

Discussion

Explicit emphasis on repetitions was an effective method of increasing their production frequency. However, even with such instructions, children produced fewer repeats than would be expected in a random set (5 out of 16 children producing more repetitions than chance predicts). Furthermore, the response change was most evident for immediate repetitions, not for longer gaps (i.e. the instructional effect was significant only for the Phi2 score).

The study yielded weak evidence for developmental improvements in performance, with correlations between age and task compliance/paired redundancy. Both impressionistic inference and formal testing suggest that children below 5 years of age may not provide a useful study group for randomization. Children aged 5 and under are likely to: (a) decline to do the task, (b) fail to produce responses quickly enough, (c) have substantial difficulties producing responses on demand (i.e. keeping to a response rhythm), and (d) merely produce repeated recitations of the numbers 1 through 10 or call on certain paired responses repeatedly.

Inevitably, the above conclusion requires caution. One may find precocious performers around this age. For example, one child of 5 years 3 months began their sequence '10, 3, 4, 1, 5, 9, 10, 10, 3, 1', appearing to show a good grasp of task requirements and an ability to carry these out. One may also be presented with 6-year-olds who do not comply with task requirements. Additional instructions, more extensive training sessions and slower response rates might enable more children at or below 5 years of age to complete the task satisfactorily. Informal observation also suggested that children might find it easier to make choices given stimulus support, that is, if the numbers being randomized were visible. This might make an ordered sequence even more prepotent, but would also remove the requirement to generate representations internally (Towse, 1998). Some of the children who failed to furnish a complete response set appeared more willing, subsequently, to choose numbers from a visually presented 'list'.

That children as young as 5 years can produce strings of numbers that are not entirely predictable, given only relatively brief instructions and a quite rapid response rate, is testament to their cognitive abilities. Thus, while they may be susceptible to failing Piagetian tests of logic (though see Donaldson, 1978), and while they appear to have limited insight into their responses (see qualitative comments above), young children *can* nonetheless demonstrate an ability to engage in tasks that require control and flexible thought.

COMBINED ANALYSIS OF SEQUENCE GENERATION

The absence of reliable age effects, particularly with respect to Expt 2 (and, to a lesser extent, Expt 3), might be taken as challenging the notion of developmental change in randomization performance. The alternative is that there may have been insufficient power in the individual experiments to reveal age-related changes. Therefore, comparable data were combined, where children produced 70 responses at the rate of one item every 2.5 s, using the numbers 1–10. To increase sample size, 49 8- and 10-year-olds were also included from a separate study (see Towse & Mclachlan, 1999). Data from the resulting corpus of 126 children were then subjected to a battery of randomization tests (see below). Nine children were then excluded owing to the presence of univariate outlier scores; with the remaining 117 children, correlational and principal-components analyses were performed.

The correlational analysis showed some highly reliable, though modest, performance changes with age (see Table 4). Chain lengths become more appropriate with age, and older children distribute response alternatives more evenly. However, repetition behaviour overall did not change with age. These findings confirm the age effects in Expt 1, while also suggesting that age improvements are in fact relatively weak. To investigate the relationship between measures of randomness further, 16 tests, as detailed in Appendix 1, were entered into a principal-components analysis. Analysis with varimax rotation led to the extraction of four factors (from inspection of the scree plot, and on the basis of eigenvalues greater than 1), together explaining 63.3% of the variance. A factor loading cut-off of .45 (Tabachnick & Fidell, 1996) was used to assay which variables substantially loaded on the factor matrix.

Table 4. Correlations between age and random sequence quality across experiments

	Age	TPI	R	Phi2	Phi7
TPI	.32*				
R	−.31*	−.01			
Phi2	−.12	−.29*	.05		
Phi7	.06	.02	.08	.04	

* $p < .001$.

TPI: chain lengths; R: distributions; Phi2: immediate repetitions; Phi7: distant repetitions.

The resulting factor structure is presented in Table 5, and factor loadings are described

Table 5. Factor structure from combined analysis of data. Factors and variables are ordered by size, factor labels are taken from Towse & Neil (1998)

Factor 1: prepotent associates	Factor 2: equality of response usage	Factor 3: long repetitions	Factor 4: short repetitions
TPI	Mean rep. gap	Phi (7 gram)	Phi (3 gram)
Runs	R	Median rep. gap	Phi (2 gram)
RNG	Coupon	Phi (6 gram)	Phi (4 gram)
A		Modal rep. gap	
RNG2			
Median rep. gap			

fully in Appendix 2. The factor structure was similar to that reported in Towse & Neil (1998) for written number generation among adults, and so the same factor labels are adopted here. Thus, analysis shows children's random generation reflects the extent of their reliance on prepotent or associated sequences, the propensity to distribute response alternatives evenly, the (lack of) repetition of responses over short sequences and repetition over somewhat longer sequences.

A number of points are noteworthy. First, analysis vindicates the use of multiple measures of randomization that reflect different performance factors. Second, analysis highlights the complexity of random generation, indicating different 'strategies' or processes making up response selections. Third, despite the necessarily exploratory nature of the analysis, it does converge with experimental findings. One might well be wary of over-interpreting the factor model given the modest sample size and the logical inability to incorporate all possible tests of randomness. Nonetheless, the results are similar to adult data (Towse & Neil, 1998) with just under 90% correspondence between the significance of factor loadings. Thus, random generation processes in children and adults appear similar in type, even where they differ in efficiency.

GENERAL DISCUSSION

Random generation tasks have been widely used among adults as an executive operation, thought to impose substantial demands insofar as participants must attempt to inhibit prepotent responses (for a review, see Brugger, 1997). Surprisingly, perhaps, the task has not been studied extensively in developmental contexts. Based on the attempt to fit production 'rules' to performance, Rabinowitz *et al.* (1989) suggested that there are developmental improvements in random generation between 7 years of age and adulthood (see also Thomas, 1969). Here, the calculation of conventional and varied randomness measures confirms modest developmental changes, when including an assessment of younger children used previously. In addition, the present studies show that children's randomization behaviour is sensitive to response speed, vocabulary range and specific instructional requests, but fairly insensitive to the general contextual framing for explaining the task. The age-related changes in performance are interesting because

Barkley (1997) has portrayed children with attention difficulties as being developmentally delayed in terms of 'inhibitory regulation'. Accordingly, they might be expected to find the avoidance of prepotent response chains particularly problematic. Meanwhile, other aspects of the task, such as repetitive behaviour, may be characteristic of other developmental disorders (see e.g. Turner, 1997).

In these studies, children's performance in any particular experimental condition was assessed on the basis of 70 responses. This was designed as a compromise between the benefits in deriving more reliable and stable randomness scores, and the advantages in maintaining children's task motivation by shortening assessment. In the main, this methodological decision appears to have been successful, since reliable and significant effects of response speed, set size and emphasis on repetitions were obtained. It is possible, however, that the relative scarcity of data contributed to the absence of age effects.

The present results are clearly consonant with the view that executive tasks incorporate a variety of cognitive processes and mechanisms. The umbrella term 'executive function' may have some utility as a reference to a class of general cognitive skills. Nonetheless, it is important to acknowledge that 'executive' tasks appear to be bound together by conceptual similarity and not necessarily by a reference to a common cognitive system. Furthermore, 'inhibitory control' may refer to rather dissociable effects; the prevalence of prepotent sequences varied across manipulations while repetition response suppression remained invariant (except under instructional manipulation in Expt 3, and even here repetitions remained underrepresented in sequences).

In conclusion, it is argued that random generation is potentially a useful cognitive tool that can be administered to children. As with other tasks, it would be misleading to regard random generation as a pure measure of some cognitive ability. Rather, it involves the recruitment of several different processes. These different processes are evident from response speed, vocabulary size and repetition avoidance phenomena in current datasets, converging with the results of principal-components analysis. In general, there is considerable overlap in performance with that found among adults, suggesting that developmental differences reflect quantitative performance change, rather than the application of different strategies or the development of a single, specific task component. Though random generation is an unusual and novel task, children's responses from coherent attempts at randomization, indicating a perhaps surprising flexibility and sophistication in their cognitive activity.

Acknowledgements

The authors are grateful to the staff and pupils of St Jude's, Thorpe, and Thorpe Lea schools for their cooperation. A computer program for measuring response characteristics in random response sequences is available from the first author on request. Comments from Alan Baddeley, Chris Jarrold, James Russell and John Wilding have been most helpful in the development of this study.

References

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: Clarendon Press.
Baddeley, A. D. (1992). Is working memory working? *Quarterly Journal of Experimental Psychology*, **44A**, 1–31.
Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, **49A**, 5–28.
Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive function: Constructing a unifying theory of ADHD. *Psychological Bulletin*, **121**, 65–94.
Bar-Hillel, M. & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics*, **12**, 428–454.
Bjorklund, D. F. (1990). *Children's strategies: Contemporary views of cognitive development*. Hillsdale, NJ: Erlbaum.
Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual and Motor Skills*, **84**, 627–661.
Dempster, F. N. (1993). Resistance to interference: Developmental changes in a basic processing mechanism. In M. L. Howe & R. Pasnak (Eds), *Emerging themes in cognitive development*, pp. 3–27. New York: Springer-Verlag.
Diamond, A., Prevor, M. B., Callender, G. & Druin, D. P. (1997). Prefontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society for Research in Child Development*, **62**, 4.
Donaldson, M. (1978). *Children's minds*. London: Fontana.
Ginsburg, N. & Karpiuk, P. (1994). Random generation: Analysis of the responses. *Perceptual and Motor Skills*, **79**, 1059–1067.
Halliday, M. S. & Hitch, G. J. (1988). Developmental applications of working memory. In G. Claxton (Ed.), *New directions in cognition*, pp. 193–222. London: Routledge and Keegan Paul.
Levin, H. S., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., Ringholz, G., Ewing-Cobbs, L & Fletcher, J.M. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology*, 7(3), 377–395.
Neuringer, A. (1986). Can people behave 'randomly'? The role of feedback. *Journal of Experimental Psychology: General*, **115**, 62–75.
Norman, D. A. & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds), *Consciousness and self-regulation*, pp. 1–18. New York: Plenum.
Rabinowitz, F. M., Dunlap, W. P., Grant, M. J. & Campione, J. C. (1989). The rules used by children and adults in attempting to generate random numbers. *Journal of Mathematical Psychology*, **33**, 227–287.
Robertson, C., Hazlewood, R. & Rawson, M. D. (1996). The effects of Parkinson's Disease on the capacity to generate information randomly. *Neuropsychologia*, **34**, 1069–1078.
Sachs, L. (1978). *Applied statistics: A handbook of techniques*. Berlin: Springer-Verlag.
Salway, A. F. S. (1991). Random number generation in the working memory dual-task paradigm. Unpublished PhD thesis, Aberdeen.
Sophian, C. (1997). Beyond competence: The significance of performance for conceptual development. *Cognitive Development*, **12**(3), 281–303.
Spence, D. P., Wolitzky, D. L. & Pezenik, J. (1969). Random generation of digits as a measure of attention: Relation to text redundancy and recall. *Journal of Verbal Learning and Verbal Behavior*, **8**, 9–15.
Tabachnick, B. G. & Fidell, L. S. (1996). *Using multivariate statistics*, 3rd ed. New York: HarperCollins.
Thomas, H. B. G. (1969). Age and preferred information rates in a randomisation task. *Journal of Child Psychology and Psychiatry*, **10**, 177–193.
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 behavior: A review of methods used and a computer program for describing performance. *Behavior Research Methods, Instruments & Computers*, 30(4), 583–591.
Towse, J. N. & Mclachlan, A. (1999). *A developmental study of random generation in children*. Technical report (CDRG6), Royal Holloway University of London (http://www.pc.rhbnc.ac.uk/papers/tr.html).
Towse, J. N. & Valentine, J. D. (1997). Random generation of numbers: A search for underlying processes. *European Journal of Cognitive Psychology*, **9**, 381–400.

Turner, M. (1997). Towards an executive dysfunction account of repetitive behaviour in autism. In J. Russell (Ed.), *Autism as an executive disorder*, pp. 57–100. Oxford: Oxford University Press.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, 77(2), 65–72.

Warren, P. A. & Morin, R. E. (1965). Random generation: Number of symbols to be randomized and time per response. *Psychonomic Science*, 3, 557–558.

Welsh, M. C., Pennington, B. F. & Groisser, D. B. (1991). A normative- developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology*, 7, 131–149.

Wiegersma, S. (1976). Response generation: Type of analysis and the limited capacity hypothesis. *Acta Psychologica*, 40, 331–342.

Received 21 August 1997; revised version received 1 October 1998

Appendix 1

Tests of randomness are described, along with computational details, more fully in Towse & Neil (1998) and Ginsburg & Karpiuk (1994). Performance in this study was scored in the following way (listed first for the primary measures, then by factor structure):

R (Redundancy). The degree to which response alternatives are equally likely to be chosen. Higher values reflect less even distributions (e.g. where a subset of all responses is used).

TPI (Turning Point Index). Assesses the number of responses that mark a change between ascending and descending sequences (points that represent local peaks and troughs in a time-series plot). Too few changes or too many changes result in values less than and greater than 100, respectively.

Phi(n). The phi index is a measure of repetition tendency over a given length n for binary sequences (or sets converted into binary streams). On the basis of known frequencies of string permutations of length n-1, Phi measures whether n-gram strings repeat or alternate end elements. A value of zero represents appropriate repetitions, negative values relative absence of repetition, positive values relative predominance of repetition. Phi is measured for lengths 2 to 7.

RNG (Random Number Generation). An index of paired redundancy, or the association between one choice and the next. Values range between 0 and 1, with higher scores indicating greater predominance of certain two-item combinations (that is, uneven distribution of paired values).

Runs. The variability in the lengths of ascending sequences throughout the response set.

A (Adjacency). The frequency of one particular stereotyped paired sequence – adjacent values on a number line.

RNG2 (Random Number Generation 2). Computational details correspond to the RNG scores, but here alternate responses are paired together (the first with the third, the second with the fourth, etc.).

RG (Repetition Gap). Measures of the mean, median and modal Repetition Gap scores, based on a table of repetition distances that reflects the number of items appearing between successive appearances of a response alternative.

Coupon. Across the entire response set, the mean number of responses emitted before the complete vocabulary has been used.

Two other measures of randomness, NSQ (Null-Score Quotient) and Phase-Length(1) frequency were not included because of the logical, computational and statistical correspondence with the above measures. This also applies to variants of measures not implemented within the analysis software, such as the Wallis-Moore phase frequency test (Sachs, 1978).

Appendix 2

Factor loading scores from principal component analysis of comparable data in Expts 1–2 and Towse & Mclachlan (1999). Values in bold denote inclusion of a variable on a particular factor.

	Factor 1	Factor 2	Factor 3	Factor 4
TPI	−.852	.138	−.025	−.227
Runs	**.831**	−.168	−.084	−.118
RNG	**.825**	.317	−.155	.096
A	**.797**	−.109	−.129	.370
RNG2	**.645**	.324	.073	−.197
Repetition gap (mean)	.080	**−.856**	.046	.043
R	.122	**.837**	.162	.082
Coupon	−.057	**.613**	.205	−.099
Phi (7 gram)	.086	−.010	**.767**	−.031
Repetition gap (median)	**−.515**	−.045	**−.698**	.197
Phi (6 gram)	.107	−.082	**.615**	.045
Repetition gap (mode)	−.377	.095	**−.568**	−.315
Phi (3 gram)	.002	.018	−.063	**.856**
Phi (2 gram)	.148	−.193	.163	**.703**
Phi (4 gram)	.079	.412	−.058	**.543**
Phi (5 gram)	−.173	.446	.336	−.017