Intelligence 86 (2021) 101543

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

String-pulling in the Greater Vasa parrot (*Coracopsis vasa*): A replication of capacity, findings of longitudinal retention, and evidence for a species-level general insight factor across five physical cognition tasks

Michael A. Woodley of Menie [a,*], Mateo Peñaherrera-Aguirre [b], Anthony M.R. Woodley [c]

[a] *Vrije Universiteit Brussel, Center Leo Apostel for Interdisciplinary Studies, Brussels, Belgium*
[b] *University of Arizona, Department of Psychology, Tucson, AZ, USA*
[c] *Independent, UK*

A B S T R A C T

Spontaneous solving of an insight-based means-end reasoning task (the string-pulling problem) is observed in an adult male captive bred Greater Vasa parrot (*Coracopsis vasa* Shaw, 1812), with an efficiency of 66%, replicating previous work in a singleton context. This case report adds to the existing literature on this species by also demonstrating longitudinal retention, specifically the same bird was found to be able to re-solve the simple form of the problem after a period of seven years (the bird was first tested in 2013, and re-tested in 2020), with an efficiency of 43% (the difference between efficiencies was not significant, $\chi^2 = 0.991$, $p = .319$). In a second analysis, species-level data across five patterned string-pulling tasks involving 14 parrot species were reanalysed, revealing that the Greater Vasa parrot exhibited the greatest general competence among those evaluated. A 'general insight factor' (GIF) was also found across taxa, the loadings onto which exhibit positive and large-magnitude associations with the correlation between fission-fusion flocking intensity and indicator level performance ($r = 0.831$), and also positive small and modest-magnitude associations with the correlation between relative brain size and indicator-level performance, and the magnitude of average pair-wise species differences in performance across indicators ($r = 0.219$ and 0.365 respectively). Finally, the theoretical implications of these findings are discussed.

1. Introduction

The string-pulling problem is a well-established ethological paradigm for identifying insight-based means-end type reasoning in animals, and in birds in particular (Jacobs and Osvath, 2015). Means-end type reasoning is a mode of problem solving which requires both *anticipation* and *implementation* of physical processes in order to attain a specific goal. Animals that engage in this form of problem solving must be able to imagine or improvise the solution (Jacobs & Osvath, 2015; Werdenich & Huber, 2006).

The simple form of the string-pulling problem involves vertically suspending a reward (typically food) on the end of a piece of string out of reach of the subject. Successfully solving this problem involves the coordination of limbs and other appendages (such as the beak in the case of birds) in preventing slippage of the string while 'reeling in' the reward (Jacobs and Osvath, 2015). More complex variants of this problem exist in which multiple strings coupled with other apparatuses can be used to examine higher-order insight, planning, and even cooperation behaviours among animals (Emery, 2006; Werdenich & Huber, 2006). The first (recorded) observation of birds solving this sort of problem was made by Pliny the Elder, who noted captive Goldfinches pulling strings with small buckets tied to their ends, so as to raise water (Jacobs & Osvath, 2015).

Thus far, string-pulling has been studied in more than 160 bird and mammal taxa, yielding more than 200 studies (Jacobs & Osvath, 2015). However, only certain species have demonstrated the capacity to solve the problem spontaneously upon being first presented with it and without prior exposure to the task, or an opportunity to otherwise socially learn the solution rules from observing its successful completion in other individuals (Audet, Ducetaz, & Lefebvre, 2016; Heinrich, 2000; Pepperberg, 2004; Werdenich & Huber, 2006). There are also debates concerning the ecological validity of apparent instances of spontaneous solving in the context of singleton trials (i.e., when the experimental subject is evaluated on its own) versus group trials (where the experimental subject is evaluated in a context where conspecifics are present). Krasheninnikova and Schneider (2014) summarise this problem as follows; "[t]esting animals individually in problem-solving tasks limits distractions of the subjects during the test, so that they can fully concentrate on the problem. However, such individual performance may not indicate the problem-solving capacity that is commonly employed in the wild when individuals are faced with a novel problem in their social groups, where the presence of a conspecific influences an individual's

---

* Corresponding author.
E-mail address: Michael.Woodley@vub.ac.be (M.A. Woodley of Menie).

https://doi.org/10.1016/j.intell.2021.101543
Received 30 June 2020; Received in revised form 8 April 2021; Accepted 10 April 2021

behaviour.” (p.1227).

Among the Psittaciformes (the order comprising the “true” parrots, New Zealand parrots, and cockatoos), thus far the capacity to solve the simple and more complex variants of this problem has been observed in 36 taxa (Jacobs & Osvath, 2015). Certain parrot species, such as the Kea (*Nestor notabilis*) and African Grey (*Psittacus erithacus*), both of which have the capacity to solve this problem, are known for being extraordinarily intelligent, and are thus popular objects of ethological study (e.g., Diamond & Bond, 1999; Pepperberg, 2002).

In the first analysis, we replicate an observation of spontaneous solving of the simple form of this problem in an unusual psittacine (a member of a subfamily within the order Psittaciformes) that has only recently attracted serious attention from ethologists by virtue of having been found to engage in spontaneous tool-use under captive conditions; the Greater Vasa parrot (*Coracopsis vasa*). Specifically, we report: 1) this taxon’s capacity to solve this problem in a singleton context (the previous report of this taxon being able to solve this task employed a group setting with 10 birds; Krasheninnikova, 2014), and 2) we demonstrate longitudinal retention of this capacity, with the same bird being able to solve the problem spontaneously after a seven-year period had elapsed. In the second analysis, we consider this species’ competence on five more complex (patterned) string-pulling tasks relative to 13 other psittacine taxa, via reanalysis of existing data. We also examine the latent structure among species differences in performance on these more complex string-pulling tasks, in addition to examining the moderating influence on the covariance among these of various (behavioral ecological and neuroanatomical) correlates of string-pulling.

1.1. *The Greater Vasa parrot*

The Greater Vasa parrot (*Coracopsis vasa*) is an Old World, long-winged psittacine, whose habitats are located in Madagascar and surrounding islands. They exhibit several unusual morphological and ethological adaptations not found in the majority of other psittacines. These include rapid maturation (relative to their size), an unusual degree of cloacal engorgement in males whilst mating (which peculiarly often involves the birds forming a copulatory tie while standing side-by-side rather than mounting one another), extensive loss of head plumage in females, along with skin yellowing and a change in plumage coloration from black to brown while in breeding condition (Arndt, 2000; Ekstrom, 2002; Ekstrom, Burke, Randrianaina, & Birkhead, 2007; Wilkinson & Birkhead, 1995), nocturnality (Arndt, 2000), an unusual ‘butterfly’ like wing-stroke, a series of very distinctive calls, and a strong preference for ground-based activities, such as foraging (Arndt, 2000; Forshaw, 2010). Another distinctive component of the species’ ethological profile is the presence of a complex cooperative polyandrogynous breeding system characterized by high levels of extra-pair copulation and alloparenting (care provided by individuals other than the younglings’ parents; Kenkel, Perkeybile, & Carter, 2017) involving groups of males feeding multiple females and their chicks (Ekstrom, 2002; Ekstrom et al., 2007; Lovegrove, Perrin, & Brown, 2011; Wilkinson, 1994). In captivity, they are also known to be moderately difficult to breed (“occasionally achieved”, Arndt, 2000, p. 158) and are also reasonably long-lived, with the record being 53.9 years (Brouwer, Jones, King, & Schifter, 2000). Famed German naturalist Alexander von Humboldt (1769–1859) had a pet Greater Vasa “Jacob”, who is known to have lived over 30 years (Becker & Fiebig, 2004).

The Greater Vasa has been found to be capable of spontaneous tool-use. These observations were made in 2015 and involved a group of six captive Greater Vasas employing stones and date pits to grind calcium from the surface of sea shells. They were also noted to have shared these tools with one another, a behaviour that may be unique to this psittacine. Observations of spontaneous tool use in parrots are exceptionally rare (Lambert, Seed, & Slocombe, 2015), and is suggestive of very high species-typical intelligence.

2. Analysis 1: experimental replication of simple string-pulling in the Greater Vasa

In terms of complex means-end reasoning tasks, such as string-pulling, Greater Vasas, along with 13 other psittacine taxa, have been studied by Krasheninnikova (2014), as part of dissertation work on the comparative psychology of physical cognition across psittacine taxa. The research involved examination of both simple and patterned forms of string-pulling, and employed a collection of ten Greater Vasas (five males and five females), which were evaluated at the Parrot Zoo in Skegness, UK. The Greater Vasas were found to be highly competent across both the simple and more complex patterned versions of the string-pulling task. Critically these Greater Vasas were evaluated in a group setting. As Krasheninnikova and Schneider (2014), have noted, there are advantages to grouping the birds for the purpose of increasing the ecological validity of observations, however singleton or individual testing *can* limit distractions of a sort that might make them less inclined to solve the task efficiently. Here we will evaluate the performance of the Greater Vasa parrot in a singleton context. We also have the ability to test for another component of string-pulling, one that to our knowledge has never been studied before, specifically longitudinal retention, which will allow us to determine whether or not the bird retains the capacity to spontaneously solve the same problem upon being presented with it after several years have elapsed. This tests for the presence of age-related changes in performance.

2.1. *Methods*

A captive bred, hand reared male Greater Vasa parrot (ring number: 68 PSUK98 W, named “Diesel”) was sourced from the collection of the lead author. The bird was 15 (as of 2013) years of age, and is housed in a moderately large cage (1.5 m × 0.8 m × 0.52 m) within the kitchen/dining area of the lead author’s parent’s house. The kitchen/dining area currently also houses a female African Grey of equivalent age, and previously also housed a male Eclectus (*Eclectus roratus*) and female Jardines Parrot (*Poicephalus gulielmi*), both deceased.

The bird was first presented with the simple form of the string-pulling task in March 2013. This involved a piece of twine, approximately 60 cm in length with a food reward (bait) suspended vertically at one end. The string was tied to the back of a chair, on which the bird could perch. The objective of the task was for the bird to obtain the bait, which the bird could see ‘dangling’ out of reach on the end of the string. The setup was employed for two rounds of experimentation, one in March 2013 (single trial) and a second in June 2020 (to determine retention, two trials, one refamiliarization and one main). The experiments were observational, non-invasive, and were conducted in accordance with best practices guidelines (Guidelines for the Use of Animals, 2018).

2.2. *Results*

In the 2013 trial the bird was able to solve this problem spontaneously upon first being presented with the task. The bird showed no aversion to the task (no neophobia) and took to solving it immediately upon being presented with it (i.e., qualitatively, the bird’s latency of approach to the task could reasonably be described as ‘very low’). No trial-and-error learning period was involved in the 2013 trial, and the bird successfully completed the trial after 24 s had elapsed. The bird was also found to be able to re-solve the problem later, over a testing interval of seven years. In the 2020 two-trial session, the solution time in the main trial[1] was approximately 17 s. An attempt was made to homogenize the testing conditions with those of the 2013 trial to the greatest extent possible (e.g., the bird was tested in the same location in the kitchen, with similar type of string and bait-type – a piece of biscuit).

Fig. 1 shows four stills of the bird taken using an iPhone 4 at various stages of solving the problem from the 2013 trial. Consistent with work on certain other psittacine taxa, such as the Galah cockatoo (*Eolophus roseicapilla*) (Krasheninnikova, 2014), the Greater Vasa was observed to primarily employ the *upright pull* strategy for reeling in the bait (as can be seen in Fig. 1). This involves pulling the string vertically, until the bird was in an upright position, then grasping the string with the foot before reaching down to pull up more string. In the 2013 trial, 12 beak-string contacts were counted, of which 10 were defined as successful as they resulted in lifting of the string, and two were defined as ineffectual (these involved pulling at loops which did not raise the string). In the 2020 main trial, other, flawed strategies were also observed to be in use. These included one instance of *flipping* which involves pulling the string over to the other side of the perch. And one instance of *side-walking*, which involves grasping the string with the beak and pulling it horizontally by walking away from the anchor-point. In the 2020 trial, the bird was observed to perform seven beak-string contacts in the course of solving the problem. Of these actions five resulted in the bait being raised closer to the bird. The two other actions resulted in the bait either not being raised, or being dropped. The footage of the two trials was independently reviewed by MPA, who noted the same numbers of each event as the lead author (inter-rater reliability = 1.00). The efficiency of the bird in each trial was calculated using the following formula (modified from Krasheninnikova, 2014, p. 38):

$$(f_s – f_u)/t = e^*100 \tag{1}$$

where f_s equals the frequency of successful actions, f_u equals the frequency of unsuccessful actions, t equals the total number of actions, and e equals the efficiency (expressed as a percentage). A composite of two pieces of footage, one 24 s piece from March 2013 and a second 17 s piece of the main June 2020 trial, taken using an iPhone 4 and 7 respectively, has been deposited as an online supplement (Supplement 1).

Table 1 presents the results for both the 2013 and 2020 trial sessions for comparison. A "yes" response to the approach item indicated that the bird was observed to engage with the task (low neophobia) and furthermore did so quickly (low latency of approach). Other relevant data presented include the numbers of trials in each session, and estimates of the mean solving time in approximate seconds. To have solved the task in this instance involved the bird successfully reeling-in and taking the bait/reward in his beak.

2.2.1. Summary of results

Although the Greater Vasa has already been studied in relation to its abilities to solve both the simple and more complex (patterned) version of the string-pulling problem, we were able to observe an example of this species solving the problem as a singleton, without other conspecifics being present. Recall that Krasheninnikova (2014) tested their Greater Vasas as part of a group of 10 birds. Singleton testing has certain advantages, as it allows the bird to focus exclusively on the task, without risk of being distracted by other birds. As Krasheninnikova and Schneider (2014) have noted (and as was discussed in the introduction), the singleton context might also risk reducing the ecological validity of apparent spontaneous string-pulling behaviour. This is because, in naturalistic settings, insight-based reasoning would likely only take place within the context of conspecific interactions. In their analysis of performance differences between Orange Winged Amazon parrots (*Amazona amazonica*) tested under both singleton and group settings, they found no difference in performance capability. However, they noted that birds were more willing to attempt the task when tested in the presence of conspecifics. We found no reluctance to engage the task in the Greater Vasa studied here (the response latency was low, with the bird 'taking to the task' very quickly) despite the singleton testing context. Of course, this observation concerns an *N* of one, so may well stem from some peculiarity associated with Diesel – rather than this being reflective of some species-typical characteristic.

When taken together with Krasheninnikova's (2014) results, it can be inferred that the Greater Vasa is apparently quite a versatile solver of string-pulling problems across different social contexts (e.g., singleton vs. group). We were also able to recover the spontaneous string-pulling problem-solving behaviour in the same bird assessed longitudinally after a period of seven years. There was a decrease in solving efficiency from 67% to 43% (it should be noted that the bird also failed to solve the 2020 refamiliarization trial), however the difference in relative efficiency between the main trials was not statistically significant at the $p \leq .05$ level ($\chi^2 = 0.991$,[2] $p = .319$). Longitudinal designs tracking age-related changes are never, to the knowledge of the authors, employed in studies of string-pulling, but could yield insights into what age-related changes may occur on these tasks (if any).

3. Analysis 2: a general insight factor across 14 psittacine taxa

Krasheninnikova (2014) presents performance data on five patterned string-pulling tasks for 14 psittacine taxa, covering three Families, the Cacatunidae, the Psittacidae, and the Psittaculidae. Employing a regression-tree analysis, Krasheninnikova (2014) was able to analyse the contributions to performance variance (as measured using the proportions of correct choices) from a variety of ecological factors, finding that fission-fusion flocking intensity (this being the tendency for groups to change in size via the fission and fusion of subunits in response to activity patterns and resource distributions; Aureli et al., 2008) accounted for the majority of the variance in performance for four of the five tasks. It should be noted that Greater Vasas are rated as 'high' on fission-fusion flocking intensity (Krasheninnikova, 2014). The fact that a *single* behavioral ecological factor was found to account for the largest portion of performance variance across so many tasks raises the possibility that diverse forms of string pulling might be reflective of an underlying general factor among the different physical cognitive domains tapped by different experimental modes of string pulling, a possibility apparently not considered by Krasheninnikova (2014).

3.1. Methods

We constructed a 'general insight factor' (GIF) of cross-species differences in string-pulling via unit-weighted factor scoring based on the species difference in proportion of correct choices (scaled 0 to 1.00) in the five patterned string-pulling tasks listed in Krasheninnikova's (2014) Supplemental Table (S1, p. 29). The five tasks were as follows:

1) *Directedness*: This measures whether the bird is food-directed or not by presenting it with two strings, one baited and the other not.

---

[1] In the refamiliarization trial, the bird did not successfully solve the problem. A major issue was with the string, which was quite a bit longer in this trial (closer to a meter). The bird was able to partially reel in the bait, but appeared to get bored, dropping the string completely altogether after about 20 s had elapsed. In the subsequent, main replication trial, the string was shortened to approximately 40 cm so as to mitigate the attention-span issue. The footage of this practice trial will be made available upon request.

[2] For the comparison of proportions test we used the number of beak-string contacts as equivalent to the sample size, so for the 2013 trial, $N = 12$, and for the 2020 trial, $N = 7$.

Fig. 1. The 2013 trial. (Top left; time index 00:00:00) The bird initially evaluates the problem. (Top right; time index 00:00:08) The bird grasps a loop of string with his beak pulling it upright, then grasps it with his foot. (Bottom left; time index 00:00:19) The bird uses a combination of beak-tugging and loop-grasping to reel in more of the string. (Bottom right; time index 00:00:24) The bird takes receipt of the bait after having successfully completed the task.

Table 1
The two sessions (2013) and (2020), along with qualitative information on the birds' approach to the trial, trial numbers per session, mean solving time per session, and efficiency (%).

Session	Did the bird approach the task?	Number of Trials	Solving time (approx. seconds to reel-in the bait)	Efficiency
March 2013 session	Yes (low latency)	1	24	67%
June 2020 session	Yes (low latency)	2 (one refamiliarization trial, and one main trial)	17 (main trial)	43%

2) *Colour*: This measures the ability to make a string colour-food connection with crossed strings one of which is coloured and the other not. The colour allows the bird to trace a path to the bait.
3) *Continuity*: This measures visual comprehension of physical continuity using two crossed strings of the same colour.
4) *Connectivity*: This measures visual recognition of physical connections between objects using two parallel baited strings, only one of which is actually connected to the bait.
5) *Inhibition*: This measures inhibitory control using two parallel strings of different lengths, the shorter of which is not connected to the bait.

Performance data on the basic string-pulling task, of the sort measured in the previous experiment, was also present in the dataset. As each species 'passed' this task, gaining a simple 'yes' response there was no variation among species on this task, therefore it was not included in the factor estimation procedure. It is worth noting that relative to other factor estimation methods, unit-weighted factor scoring circumvents many of the issues related to sample specificity such as when dealing with a small number of cases and/or variables (for a review of relevant literature on the high generalizability of unit-weighted factors relative to other forms of factor estimation techniques, see: Gorsuch, 1983).

The unit-weighted factor scores are calculated as an average across the standardized patterned string-pulling test scores. Part-whole correlations between the standardized factor scores (*zf*) and the standardized indicator values (*zi*) are equivalent to traditional factor loadings ($zi^*zf = \lambda$). Prior to estimating the unit-weighted factor scores a precautionary analysis was also performed on each indicator to determine whether there was a detectable phylogenetic signal, using the package *phytools* (Revell, 2012). The presence of such a signal would indicate that special phylogenetic protocols would have to be used in estimating statistical

Table 2
Phylogenetic signal (Pagel's λ) associated with each indicator, along with *p*-value and *N* of species.

Indicator	Phylogenetic signal (Pagel's λ)	*p-Value*	*N* (species)
Directedness	5.318×10^{-5}	1.000	14
Colour	0.295	0.600	14
Continuity	0.109	0.892	14
Connectivity	1.211	0.056	13
Inhibition	5.255×10^{-5}	1.000	8

Note: It should be noted that the species numbers were small in all cases (ranging from 14 to 8 species per indicator). The results of phylogenetic signal analyses are generally considered less stable than when the *N* of species is ≥20 (Münkemüller et al., 2012; Peñaherrera-Aguirre & Fernandes, 2021).

parameters (Nunn, 2011), including factor loadings. The results of this analysis are presented in Table 2.

It was found that none of the indicators were associated with a statistically significant phylogenetic signal (in all cases $p > .05$), therefore conventional statistical approaches can be utilized in analysing these data.

In addition to the unit-weighted factor loadings, also presented in Table 3 are the correlations between species differences in fission-fusion flocking intensity (scaled as low = 1, medium = 2, and high = 3) and indicator scores, and also relative brain size and indicator scores (the variables used in estimating these correlations were both sourced from Krasheninnikova, 2014, Table S2, pp. 30–31). The *d*-values of average pair-wise species differences in performance means for each indicator were also computed. It should be noted that these values are simply difference scores and are not divided by their standard deviations as in the case of Cohen's *D*. These variables were computed for use in moderator analyses, in order to determine whether they are associated with the magnitude of the GIF loadings.

Both relative brain size variation and the intensity of fission fusion flocking might be important sources of the positive manifold among these indicators. It has been argued that brain volume might be an important source of general constrains on the total amount of cortical substrate available for investment into problem solving. Thus, larger brains contain generally more cortical substrate allowing for generally more efficient solving of diverse problems. This has been posited as a potential source of the *g*-factor at the individual differences level and possibly also between-species (this being the *processing volume theory*; Woodley of Menie, te Nijenhuis, Fernandes, & Metzen, 2016). If this is the case, then we would expect that the GIF loadings should be strongly and positively correlated with the degree to which the indicators correlate with species-differences in relative brain size. As was noted in the previous section, Krasheninnikova (2014) found indications that intensity of fission-fusion flocking was a major predictor of species-differences in performance on four of the five indicators. This is suggestive of a role for this aspect of behavioral ecology in generating covariance among indicators. If this is the case, then we would expect there to be a strong and positive correlation between the indicator GIF loading and the strength of the correlation between fission-fusion flocking intensity and species-differences in performance across indicators.

A final question is to what extent does the loading of the GIF onto its component indicators positively predict the magnitude of the species differences across indicators? The presence of positive moderation would increase the prospective validity of the factor as it would indicate that the factor is a major source of discriminant validity among species across indicators. This has been observed in analysis of species-differences in Big *G* (this being the species-level equivalent of the little-*g* of individual differences in cognitive functioning) indicator scores and factor loadings derived from other taxa, specifically primates, where it was found that the average pair-wise difference in indicator scores across 69 species was strongly and positively correlated with the Big *G* loadings associated with each indicator ($r = 0.985$; Fernandes, Woodley, & te Nijenhuis, 2014).

3.2. Results

A clear GIF emerges across the different domains tapped by the different types of patterned string-pulling task, accounting for 63.5% of the variance across species. Of all species examined, the Greater Vasa exhibited the highest GIF score (the species score averaged across the five tasks was 0.82, meaning that the Greater Vasas made the correct string choice, when presented with alternative choices, 82% of the time). The standardized GIF scores are plotted out by Family (Cacatuidae, Psittacidae, and Psittaculidae) in Fig. 2.

The Family exhibiting the highest mean of the GIF is the Psittaculidae. This Family contains Old World parrots, including the Greater Vasa. It should be noted that none of the Family-level differences were statistically significant (at $p \leq .05$) given the *N* of species per Family.

We generated a phylogenetic tree for the 14 species of psittacines consistent with the information provided by Chen et al. (2019) and Krasheninnikova (2014). After reconstructing the underlying phylogeny, we computed an ancestral character reconstruction (with maximum likelihood estimation) to determine the qualitative evolutionary trajectory of the GIF across the 14 parrot species (Fig. 3). This procedure (conducted using *phytools*, Revell, 2012) allows researchers to determine whether a trait changed in magnitude over time and to reconstruct the likely phenotypic value of the ancestral species (i.e., "statistical paleontology," Pagel, 1997; Nunn, 2011). As evidenced by the ancestral character reconstruction analysis, high GIF scores evolved at least two times in psittacines; in the Greater Vasa and also in the Spectacled Parrotlet (*Forpus conspicillatus*). In contrast, the remaining species' *z*-scores fell within a range of –1.0 to +1.0 standard deviations from the mean.

The results of the vector correlation analyses are presented in Fig. 4. The vector correlation between *r*(Rel. brain size*Indicator) and GIF λ was found to be 0.813, which is *very large* in magnitude (i.e., $r \geq 0.70$; Rosenthal, 1996). The vector correlation between *r*(FFF*Indicator) and GIF λ was found to be 0.219, which is *small* in magnitude (i.e., *r* between 0.10 and 0.29; Rosenthal, 1996). The vector correlation between the mean pair-wise species differences (*d*) associated with each indicator and GIF λ was found to be 0.365, which is *medium* in magnitude (i.e., *r* between 0.30 and 0.49; Rosenthal, 1996), None of these effect sizes were statistically significant (at $p \leq .05$) given the small an *N* of (five) indicators.

4. Discussion

The ability to solve the simple string-pulling problem was found to replicate in an individual, male 15-year-old Greater Vasa parrot. The bird performed the task in a singleton context, which might indicate that this problem-solving ability in this species generalizes outside of a group assessment setting (which is the other context in which this species has been assessed on string-pulling problems; Krasheninnikova, 2014). This is consistent with Krasheninnikova and Schneider's (2014) observations on other psittacine species, where it has been found that testing context does not seem to affect task performance.

The bird was qualitatively found to be inquisitive (low latency in approaching the task), which is consistent with Krasheninnikova's (2014) observation that relative to other psittacine species, the Greater Vasa exhibited both low mean and inter-individual variance in approach latency when presented with the basic form of the string-pulling problem (Fig. 1, p.75), suggesting heightened curiosity and venturesomeness relative to the other psittacine species tested.[3] The same bird was retested seven years after the initial testing period (at age 22), and was found to be able to spontaneously re-solve the problem, when retested (after a single refamiliarization trial), without any statistically significant longitudinal decrement in relative solving efficiency.

Factor analysis of the species difference in proportions of correct responses across five different patterned string-pulling tasks yields clear indications of a GIF, with the Greater Vasa being the most able of 14 species analysed, and the family Psittaculidae, which contains the

---

[3] This is anecdotally substantiated by the lead author who has noted that Diesel is extraordinarily curious, and will eagerly investigate any 'anomaly' in his environment, such as an open cupboard, or a new object in the room. Interestingly, despite being highly curious about physical novelty in his environment, he exhibits the opposite tendency, i.e., heightened neophobia to the presence of new people in his environment. Although approach times were not measured, it was noted that Diesel was extremely interested in the task as it was being set up, flying over to the chair to 'investigate' (interfere) while the string was being anchored.

Table 3
Unit-weighted factor loadings (λ) of a General Insight Factor (GIF) onto its five constituent indicators, along with 95% CIs, p-values, the correlations between indicator scores and fission fusion flocking intensity (FFF) and relative brain size, and the mean pair-wise difference (d) among all species for each indicator. $N =$ the numbers of species per indicator.

Indicator	λ (95% C.I.)	p-Value	r (FFF*indicator)	r (Relative brain size*indicator)	d-Value	N (species)
Directedness	0.727 (0.320, 0.908)	0.0032	0.163	−0.167	−0.045	14
Colour	0.784 (0.434, 0.928)	0.0009	0.399	0.187	0.004	14
Continuity	0.827 (0.528, 0.944)	0.0003	0.338	0.210	0.061	14
Connectivity	0.775 (0.391, 0.929)	0.0019	0.444	0.318	0.198	13
Inhibition	0.864 (0.407, 0.975)	0.0001	0.667	−0.008	0.090	8

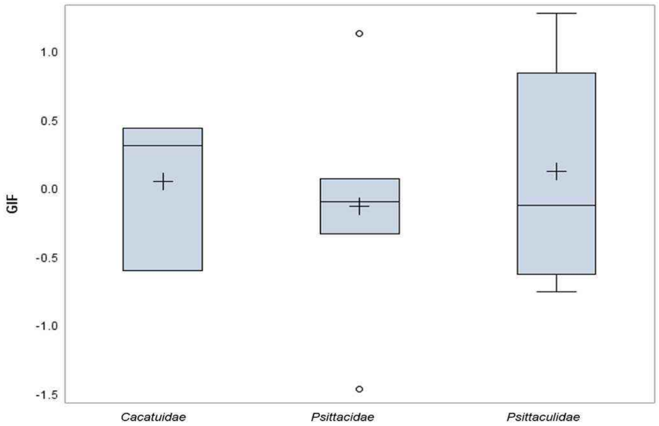

Fig. 2. Box-plots of GIF scores disaggregated by Family.

Greater Vasa parrot, exhibiting the highest mean value among the three families examined (although none of the family-level differences were statistically significant at $p \leq .05$). Ancestral character reconstruction of the GIF across the 14 parrot species revealed that high GIF has arisen twice, once in the Greater Vasa, and also in the Spectacled Parrotlet.

We initially anticipated the existence of a GIF on the basis of Krasheninnikova's (2014) finding that fission-fusion flocking intensity accounted for the largest portion of between-species variance in four of the five tasks. To comprehensively test the idea that this aspect of psittacine behavioral ecology is an important correlate of covariance among the indicators comprising the GIF, we computed a vector correlation between i) the correlation between fission-fusion flocking intensity and each indicator, and ii) the GIF λ values for each of those indicators. This yielded a very-large magnitude (but non-significant, given only five observations) correlation (0.831). The results of this analysis reinforces Krasheninnikova's (2014) suggestion that components of the social environment of psittacines might be an important source of selection pressures operating on physical cognition. Our own results suggest that fission-fusion flocking intensity is especially important, as it might also be a source of apparent species-level performance covariance among different string-pulling tasks, which might tap different, narrower facets of physical cognition within the broader domain of insight-based problem solving (such as colour discrimination, physical continuity, capacity for inhibition etc). This speculation is in line with Aureli et al. (2008), who proposed that where fission-fusion flocking is intense, there exists selection pressures favouring enhanced general information processing in response to the need to track high variability in flock composition. As fission-fusion flocking is intense among Greater Vasas (Krasheninnikova, 2014), this aspect of their behavioral ecology might have played a major role in driving their high level of GIF.

We found indications of a positive, but small magnitude (and non-significant) correlation between the vector of the correlation between relative brain size and each indicator and the vector of GIF indicator λ values ($r = 0.219$). This finding suggests that the loading of each indicator onto the GIF is only weakly moderating the strength of the indicator*relative brain size correlation. This suggests that constrains on the development of different abilities stemming from species differences in processing volume are not a major source of the covariance among the indicators comprising the GIF. This finding is consistent with human research, where it has been found in meta-analysis that the brain volume*ability correlation vector is only weaky, positively correlated with the tests g-loading vector across IQ batteries ($\rho = 0.07$, $K = 4$, $N = 246$; Woodley of Menie et al., 2016). In comparative analysis of the

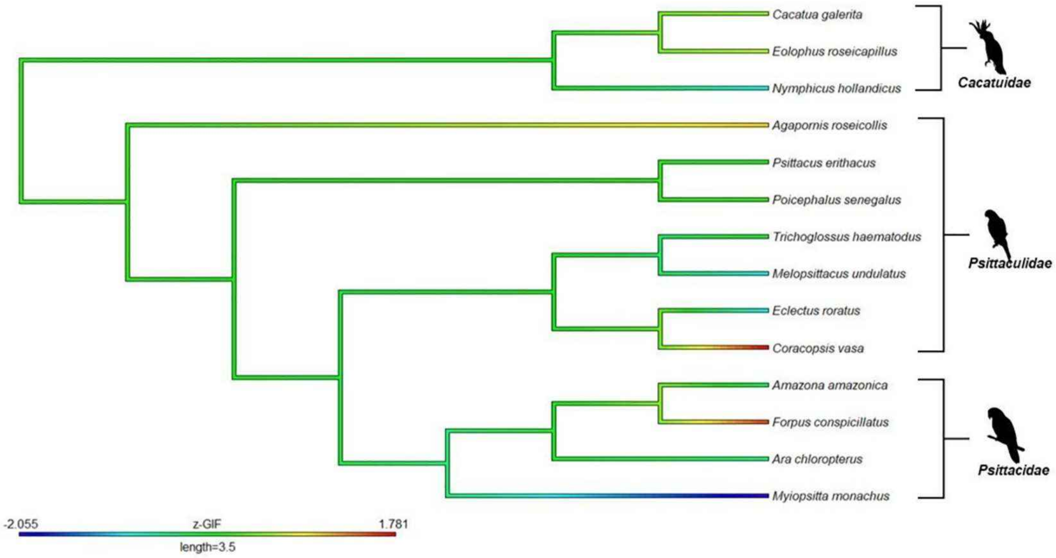

Fig. 3. Ancestral character reconstruction 'heat map' of level of GIF across 14 parrot species.

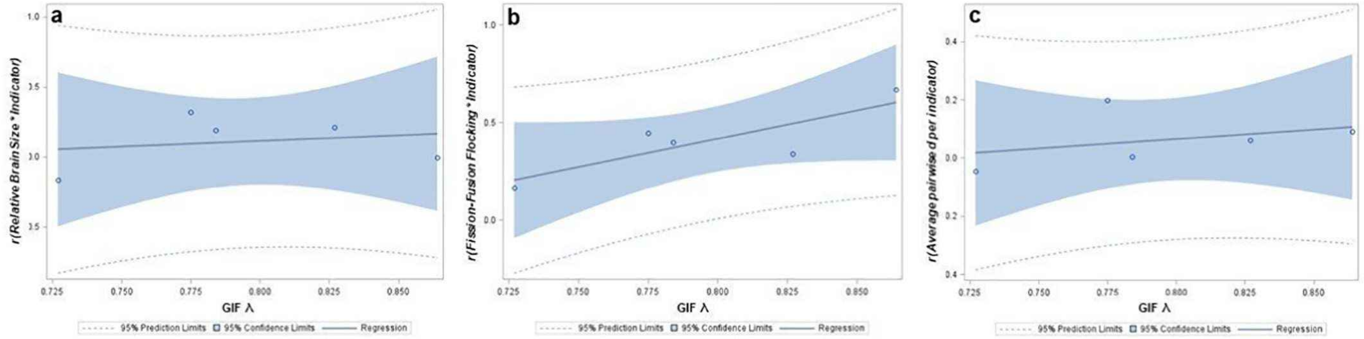

Fig. 4. Scatter plots of the vector correlations between GIF λ and *r*(Rel. brain size*indicator) (A), *r*(FFF*indicator) (B), and mean species pair-wise differences (*d*) for each indicator (C).

phylogenetic correlates of Big *G* across primate species, it has been found that both absolute and relative brain volume are poor phylogenetic correlates of Big *G* variance, with the former exhibiting much lower evolutionary lability than the latter (Fernandes, Peñaherrera-Aguirre, Woodley of Menie, & Figueredo, 2020). Only certain neuroanatomical measures, such as cerebellum size, come close to proxying the evolutionary lability of Big *G* (Fernandes et al., 2020). Although we are unable to directly examine the phylogenetic correlations between relative brain size and the GIF in Krasheninnikova's (2014) dataset (as these are available for too few species), the finding that the former's association with indicator-level performance is not strongly moderated by GIF λ, is nevertheless suggestive of the possibility that constraints arising from processing volume availability are not (primarily) responsible for driving the indicators of the GIF into correlation with one another. This in turn leads us to predict that, like primates, the more evolutionarily relevant associations are going to be between the GIF and specific avian neuroanatomical volume measures.[4] This can be tested in future research with a larger dataset.

We found indications of a positive medium magnitude, but not significant (again given the small number of observations), association between the mean pair-wise difference among species on each indicator and the degree to which the common factor loads onto each indicator ($r = 0.365$). This partially replicates Fernandes et al.'s (2014) results involving primates, especially in terms of the direction of the effect, suggesting a positive moderation effect, but not one that is nearly as pronounced as the one found across primates.

Whether the GIF can be taken as being demonstrative of the existence of high species-typical Big *G* more broadly depends on whether measures of string-pulling correlate with other, distinct (physical and non-physical) cognition measures between species, or whether they correspond to a specific, uncorrelated domain of avian cognition. The observation that the Greater Vasa is apparently an excellent 'string puller' (relative to 13 other psittacines) in addition to being one of only a tiny number of psittacines known to be capable of spontaneous tool use, and perhaps the only psittacine known to spontaneously share tools (Lambert et al., 2015) might hint at this possibility however, as enhanced ability in one domain (insight-based reasoning), seems in the case of the Greater Vasa, to go with enhanced ability in another, distinct domain of physical cognition (tool manufacture) and possibly also to a domain of social-physical cognition (tool sharing).

A reasonable and conservative interpretation of the identity of the GIF is that it likely corresponds to a broader domain of physical cognition, which is going to occupy a lower-rung in a hierarchy relative to an even broader Big *G* factor, that will draw on other cognitive domains across many different avian species. This situation is not unlike the Cattell-Horn-Carroll (CHC) model of intelligence in humans (Carroll, 1993), where there exists a general intelligence (*g*) factor at the apex of the hierarchy (Stratum III), below which are eight group factors (such as fluid and crystallised abilities; Stratum II), and below these are around 70 narrow competencies and specific skills (such as perceptual and psychomotor skills; Stratum I). In Fig. 5, we present a basic hypothetical model of what this avian Big *G* might look like in terms of latent structure across three strata.

For simplicity, we have populated Stratum 2 with only two group factors (GIF and General Tool Use; GTU). There are likely to exist other distinct co-equal group factors also (such as non-physical forms of social cognition).

The lack of strict factorial equivalence between the GIF and the hypothetical avian Big *G*, might account for the relatively weaker moderating influence of the GIF indicator λ values on the average pair-wise *d*-values by indicator noted here, compared with Fernandes et al. (2014, r=.365 vs. .985, N=5 indicators in both cases). This could be a consequence of the Big *G* factor employed in Fernandes et al. (2014) being much closer to the hypothetical Big *G* of avian cognition, than to the GIF, in terms of its respective level of latency. The presence of more domain specific variance associated with the GIF might therefore attenuate the positive moderating effect of indicator λ values on the pair-wise *d*-values. This could conceivably also be confounding the associations with relative brain volume. Estimating a 'true' Big *G* of avian cognition, encompassing multiple domains of social and physical cognition could allow for this possibility to be explored further.

Another interesting question is whether or not there exists an individual-differences *g*-factor within the Greater Vasa. Evaluation of this might be possible using the Primate Cognition Test Battery (PCTB), which has recently been adapted for use on psittacines (Krasheninnikova, Berardi, Lind, O'Neill, & von Bayern, 2019). Individual differences on string-pulling tasks have also been observed in psittacines (Cussen, 2017; Krasheninnikova, 2014). Perhaps combining simple string-pulling and its more taxing patterned variants with the domains tapped by the adapted PCTB as part of a broader individual-differences assessment protocol might yield clues as to whether this faculty belongs in the species' individual-differences *g*-matrix or not.

The fact that *N* equals one subject for our replication of string-pulling in the Greater Vasa represents an undoubted limitation. Despite this, such observations can nevertheless have ethological value. Historically a number of studies involving psittacine cognition focussed on establishing species-level proof of capacity using single birds (for a review of these, see: Cussen, 2017, a very famous example of this concerns the avian cognition research of Irene Pepperberg, much of which was conducted using a single African Grey parrot called Alex; Pepperberg, 2002). Moreover, single observations of behaviour can be usefully incorporated into comparative psychological meta-analytical databases

---

[4] We note that Krasheninnikova (2014, Table S2, p. 30) also reports species-level data for one neuroanatomical measure, specifically telencephalon volume, however these were only available for seven species. On that basis we did not consider these suitable for moderation analysis.

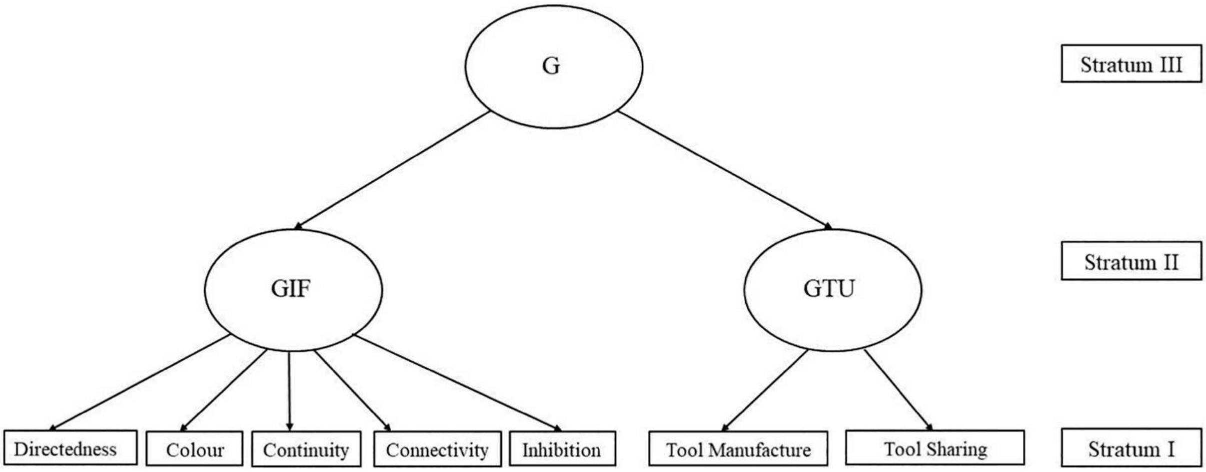

Fig. 5. Hypothetical hierarchical factor structure of Big *G* across avian species (Stratum III), with GIF and the General Tool Use (GTU) factors occupying Stratum II, and specific constituent abilities occupying Stratum I.

for the purposes of ascertaining differences between species in terms of both the levels of and covariation among various behavioral and cognitive characteristics. One confounding factor that should be considered in future experimental work is that the trials in the first analysis were conducted in the presence of the experimenter and family members. Insofar as this Greater Vasa was imprinted on humans (i.e., had been hand raised by humans from an early age and has thus developed a 'human' species identity), the presence of 'other' humans in his environment may have provided similar reinforcement cues to the presence of other Greater Vasas in the case of non-imprinted birds. Thus, future research should employ properly controlled experimental conditions, in which singleton birds are isolated for the purposes of experimental observation. Future research in string-pulling and other manifestations of avian cognition should also pay attention to longitudinal effects in e.g., efficiency.

More broadly and hopefully, this research will serve to draw the attention of those researching avian cognition to the Greater Vasa parrot, as research is increasingly converging on the observation that it is a most cognitively remarkable species worthy of substantially greater study.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.intell.2021.101543.

Declaration of Competing Interest

None.

References

Arndt, T. (2000). *Lexicon of parrots* (Vol. 4). Bretten, Germany: Arndt-Verlag.

Audet, J.-N., Ducetaz, S., & Lefebvre, L. (2016). Bajan birds pull strings: Two wild Antillean species enter the select club of string-pullers. *PLoS One, 11*, Article e0156112.

Aureli, F., Schaffner, C. M., Boesch, C., Bearder, S. K., Call, J., Chapman, C. A., … Henzi, S. P. (2008). Fission-fusion dynamics. *Current Anthropology, 49*, 627–654.

Becker, R., & Fiebig, J. (2004). Alexander von Humboldt's Vasa parrot *Coracopsis vasa* in the collection of the Museum of Natural History, Berlin. *Zoosystematics & Evolution, 80*, 283–288.

Brouwer, K., Jones, M. L., King, C. E., & Schifter, H. (2000). Longevity records for Psittaciformes in captivity. *International Zoo Yearbook, 37*, 299–316.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Chen, Y. X., Huang, Y. L., Liu, J. Q., Zhou, Y. W., & Hou, S. L. (2019). Characterization and phylogenetic relationship of the complete mitochondrial genome of Black-cheeked Lovebird, Agapornis Nigrigenis. *Mitochondrial DNA Part B, 4*, 3589–3590.

Cussen, V. A. (2017). Psittacine cognition: Individual differences and sources of variation. *Behavioral Processes, 134*, 103–109.

Diamond, J., & Bond, A. B. (1999). *Kea, bird of paradox: The evolution and behavior of a New Zealand parrot*. California, CA: University of California Press.

Ekstrom, J. M. M. (2002). *The breeding biology and behavior of the Greater Vasa Parrot Coracopsis vasa*. Doctoral dissertation. The University of Sheffield.

Ekstrom, J. M. M., Burke, T., Randrianaina, L., & Birkhead, T. R. (2007). Unusual sex roles in a highly promiscuous parrot: The Greater Vasa parrot *Caracopsis vasa*. *Ibis, 149*, 313–320.

Emery, N. J. (2006). Cognitive ornithology: The evolution of avian intelligence. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 361*, 23–43.

Fernandes, H. B. F., Peñaherrera-Aguirre, M., Woodley of Menie, M. A., & Figueredo, A. J. (2020). Macroevolutionary patterns and selection modes for general intelligence (*G*) and for commonly used neuroanatomical volume measures in primates. *Intelligence, 80*, 101456.

Fernandes, H. B. F., Woodley, M. A., & te Nijenhuis, J. (2014). Differences in cognitive abilities among primates are concentrated on *G*: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence, 46*, 311–322.

Forshaw, J. M. (2010). *Parrots of the world*. Princeton, NJ: Princeton University Press.

Gorsuch, R. L. (1983). *Factor analysis*. Hillside, NJ: L. Earlbaum Associates.

Guidelines for the Use of Animals. (2018). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour, 135*, I–X.

Heinrich, B. (2000). Testing insight in ravens. In C. M. Heyes, & L. Huber (Eds.), *The evolution of cognition* (pp. 289–305). Cambridge, MA: MIT Press.

Jacobs, I. F., & Osvath, M. (2015). The string-pulling paradigm in comparative psychology. *Journal of Comparative Psychology, 129*, 89–120.

Kenkel, W. M., Perkeybile, A. M., & Carter, C. S. (2017). The neurobiological causes and effects of alloparenting. *Developmental Neurobiology, 77*, 214–232.

Krasheninnikova, A. (2014). *Physical cognition in parrots: A comparative approach*. Doctoral dissertation. Hamburg, Germany: University of Hamburg.

Krasheninnikova, A., Berardi, R., Lind, M.-A., O'Neill, L., & von Bayern, A. M. P. (2019). Primate cognition test battery in parrots. *Behavior, 156*, 721–761.

Krasheninnikova, A., & Schneider, J. M. (2014). Testing problem-solving capacities: Differences between individual testing and social group setting. *Animal Cognition, 17*, 1227–1232.

Lambert, M. L., Seed, A. M., & Slocombe, K. E. (2015). A novel form of spontaneous tool use displayed by several captive greater vasa parrots (*Coracopsis vasa*). *Biology Letters, 11*, 20150861.

Lovegrove, B. G., Perrin, M. R., & Brown, M. (2011). The allometry of parrot BMR: Seasonal data for the Greater Vasa Parrot, *Coracopsis vasa*, from Madagascar. *Journal of Comparative Physiology B, 181*, 1075–1087.

Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K., & Thuiller, W. (2012). How to measure and test phylogenetic signal. *Methods in Ecology & Evolution, 3*, 743–756.

Nunn, C. L. (2011). *The comparative approach in evolutionary anthropology and biology*. Chicago, IL: University of Chicago Press.

Pagel, M. (1997). Inferring evolutionary processes from phylogenies. *Zoologica Scripta, 26*, 331–348.

Peñaherrera-Aguirre, M., & Fernandes, H. B. F. (2021). Phylogenetic analysis with comparative methods. In T. Shakelford, & V. Weeks-Shakelford (Eds.), *Encyclopedia of evolutionary psychological science*. https://doi.org/10.1007/978-3-319-16999-6_3605-1.

Pepperberg, I. M. (2002). *The Alex studies: Cognitive and communicative abilities of Grey parrots*. Cambridge, MA: Harvard University Press.

Pepperberg, I. M. (2004). "Insightful" string-pulling in Grey parrots (*Psittacus erithacus*) is affected by vocal competence. *Animal Cognition, 7*, 263–266.

Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology & Evolution, 3*, 217–223.

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research, 21*, 37–59.

Werdenich, D., & Huber, L. (2006). A case of quick problem solving in birds: String pulling in keas, *Nestor notabilis*. *Animal Behavior, 17*, 855–863.

Wilkinson, R. (1994). Vasa parrot's fascinating breeding behavior. *Psittascene, 6*, 9.

Wilkinson, R., & Birkhead, T. R. (1995). Copulation behavior in the Vasa parrots Coracopsis vasa and C. nigra. *Ibis, 137*, 117–119.

Woodley of Menie, M. A., te Nijenhuis, J., Fernandes, H. B. F., & Metzen, D. (2016). Small to medium magnitude Jensen effects on brain volume: A meta-analytic test of the processing volume theory of general intelligence. *Learning & Individual Differences, 51*, 215–219.