

Does bumblebee preference of continuous over interrupted strings in string-pulling tasks indicate means-end comprehension?

Reviewed Preprint

v2 • August 28, 2024

Revised by authors

Reviewed Preprint

v1 • June 5, 2024

Chao Wen , Yuyi Lu, Cwyn Solvi, Shunping Dong, Cai Wang, Xiujuan Wen, Haijun Xiao, Shikui Dong, Junbao Wen, Fei Peng , Lars Chittka 

School of Grassland Science, Beijing Forestry University, Beijing, China • Biological and Experimental Psychology, School of Biological and Behavioural Sciences, Queen Mary University of London, London, United Kingdom • Department of Psychology, School of Public Health, Southern Medical University, Guangzhou, China • Guangdong-Hong Kong-Macao Greater Bay Area Center for Brain Science and Brain-Inspired Intelligence, Southern Medical University, Guangzhou, China • Beijing Key Laboratory for Forest Pest Control, Beijing Forestry University, Beijing, China • College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou, China

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Abstract

Bumblebees (*Bombus terrestris*) have been shown to engage in string-pulling behavior to access rewards. The objective of this study was to elucidate whether bumblebees display means-end comprehension in a string-pulling task. We presented bumblebees with two options: one where a string is connected to an artificial flower containing a reward and the other presenting an interrupted string. Bumblebees displayed a consistent preference for pulling connected strings over interrupted ones after training with a stepwise pulling technique. When exposed to novel string colors, bees continued to exhibit a bias towards pulling the connected string. This suggests that bumblebees engage in featural generalization of the visual display of the string connected to the flower in this task. If the view of the string connected to the flower was restricted during the training phase, the proportion of bumblebees choosing the connected strings significantly decreased. Similarly, when the bumblebees were confronted with coiled connected strings during the testing phase, they failed to identify and reject the interrupted strings. This finding underscores the significance of visual consistency in enabling the bumblebees to perform the task successfully. Our results suggest that bumblebees' ability to distinguish between continuous strings and interrupted strings relies on a combination of image matching and associative learning, rather than means-end understanding. These insights contribute to a deeper understanding of the cognitive processes employed by bumblebees when tackling complex spatial tasks.

eLife assessment

This study provides **valuable** new insights into insect cognition and problem-solving in bumblebees. The authors present **convincing** evidence that bumblebees lack causal understanding in a string-pulling task, and find support for bumblebees instead using image-matching for this task.

<https://doi.org/10.7554/eLife.97018.2.sa2>

Introduction

String pulling is one of the most extensively used approaches in comparative psychology to evaluate the understanding of causal relationships (Jacobs & Osvath, 2015 [↗](#)), with most research focused on mammals and birds, where a food item is visible to the animal but accessible only by pulling on a string attached to the reward (Taylor *et al.*, 2010 [↗](#); Range *et al.*, 2012 [↗](#); Jacobs & Osvath, 2015 [↗](#); Wakonig *et al.*, 2021 [↗](#)). A fundamental challenge in animal cognition research revolves around unraveling the strategies that animals employ when confronted with specific tasks (de Waal & Ferrari, 2010 [↗](#); Shettleworth, 2010 [↗](#); Chittka *et al.*, 2012 [↗](#)). The complexity of the string-pulling paradigm can be altered by manipulating the number and mutual positions of the strings and reward, allowing the investigation of different aspects of cognition (Jacobs & Osvath, 2015 [↗](#); Wang *et al.*, 2019 [↗](#)). Multiple mechanisms can be involved in the string-pulling task, including the proximity principle, perceptual feedback and means-end understanding (Taylor *et al.*, 2012 [↗](#); Wasserman *et al.*, 2013 [↗](#); Jacobs & Osvath, 2015 [↗](#); Wang *et al.*, 2021 [↗](#)). The principle of proximity refers to animals preferring to pull the reward that is closest to them (Jacobs & Osvath, 2015 [↗](#)). Taylor *et al.* (2012) [↗](#) proposed that the success of New Caledonian crows in string-pulling tasks is based on a perceptual-motor feedback loop, where the reward gradually moves closer to the animal as they pull the strings. If the visual signal of the reward approaching is restricted, crows with no prior string-pulling experience are unable to solve the broken string task (Taylor *et al.*, 2012 [↗](#)).

Means-end understanding is expressed as goal-directed behavior, which involves the deliberate and planned execution of a sequence of steps to achieve a goal (Jacobs & Osvath, 2015 [↗](#); Ortiz *et al.*, 2019 [↗](#)). String-pulling studies have directly tested means-end comprehension in various species (Riemer *et al.*, 2013 [↗](#); Jacobs & Osvath, 2015 [↗](#)). In these studies, organisms are presented with two or more strings, where one is connected to a reward and the other one is interrupted; pulling the connected string indicates that animals comprehend that a continuous string is a mean to the end of obtaining the reward (Piaget, 1953 [↗](#); Wasserman *et al.*, 2013 [↗](#); Jacobs & Osvath, 2015 [↗](#); Hofmann *et al.*, 2016 [↗](#); Wang *et al.*, 2019 [↗](#)). Most animals fail in such string-pulling tasks when they have to spontaneously solve them, but they can be trained to recognize that an interrupted string is useless through trial-and-error learning (Mayer *et al.*, 2014 [↗](#); Ortiz *et al.*, 2019 [↗](#); Wang *et al.*, 2019 [↗](#)).

To our knowledge, bumblebees are the only reported invertebrates that have been trained to learn to pull a string to obtain rewards (Alem *et al.*, 2016 [↗](#); Zhou *et al.*, 2024 [↗](#)). However, the performance of bumblebees could be explained by associative learning (Alem *et al.*, 2016 [↗](#)). It remains unknown whether the bumblebees understand the connectivity of the string. The present study aims to explore the question of means-end comprehension in bumblebees when tackling string-pulling tasks. We conducted nine horizontal string-pulling experiments. These experiments were designed to manipulate factors such as string color and spatial arrangement during both the training and testing phases. Firstly, bumblebees without string-pulling experience were tested to

discriminate between strings connected to a target containing the reward and disconnected strings. In another set of experiments, we examined whether bumblebees with string-pulling experience would discriminate between connected and disconnected strings. Furthermore, we changed the color of strings in training to determine whether bees generalize features learned to solve tasks with different colored strings. In two other experiments, black tape was used to cover the strings, preventing the bees from seeing the string connected to the flower from above the table during the training phase. In one of these experiments, a green table was placed behind where the bee pulled to prevent the bee from seeing the string connected to the flower even after having reached the reward during the training phase. Finally, to further verify whether bumblebees choose strings through image learning, the straight strings were changed to coiled so that the image of the strings is distinct from those in training.

Results

In Experiment 1, we evaluated whether uninformed bumblebees with no string-pulling experience could discriminate between connected and disconnected strings. Bees were trained to retrieve a yellow flower (without a string) containing sugar water from under a transparent table (**Figure 1** [↗](#); S1 video). In the test, bees had two different flowers to choose from, one with a connected string and the other with a disconnected string (**Figure 2** [↗](#)). The strings and flowers used were discarded after each test to prevent the use of chemosensory cues. We found that 10 of 21 bees (48%, $\chi^2 = 0.05$, $P = 0.83$) chose the connected string as their first choice in the test (**Table 1** [↗](#)). Over the entire duration of the test, bees showed no preference for either string during the test [$n = 21$, generalized linear mixed model (GLMM): 95% CI = -0.09 (-0.48 to 0.29), $Z = -0.48$, $P = 0.63$; **Figure 3** [↗](#); *Table S1*], suggesting that there is no spontaneous comprehension of the significance of the gap, and that any preference for connected strings would have to be acquired through training.

We next examined whether bumblebees with string-pulling experience would discriminate between connected and disconnected strings. In Experiment 2, bees were first trained to retrieve yellow flowers from under a transparent table by pulling an attached white string (S2 video). After training, bees were presented with two different flowers: one was connected to a string and the other had a disconnected string. We found that 13 of 18 bumblebees (72%, $\chi^2 = 3.56$, $P = 0.06$) chose the connected string as their first choice in the test (**Table 1** [↗](#)). Over the entire duration of the test, bees selected the connected strings ($76 \pm 4\%$) significantly more than the disconnected strings [$n = 18$, GLMM: 95% CI = 1.06 (0.78 to 1.33), $Z = 7.56$, $P < 0.001$; **Figure 3** [↗](#), S3 video]. In addition, bees spent much more time attempting to pull the connected strings (94.67 ± 13.19 s) than the disconnected strings (13.61 ± 4.59 s) [$n = 18$, GLMM: 95% CI = 0.7 (0.02 to 0.11), $t = 2.90$, $P < 0.01$; **Figure 4** [↗](#)].

In Experiment 3, we trained another group of bees with the same procedure as in Experiment 2, but the disconnected string pointed to the flower in the test (there was no lateral displacement between the disconnected string and flower in the test; **Figure 1** [↗](#)). Again, we found that most bees (17/18, 94%, $\chi^2 = 14.22$, $P < 0.001$) choose the connected string as their first choice (**Table 1** [↗](#)). Over the course of the entire test, the percentage of the bees pulling connected strings ($79 \pm 4\%$) was significantly above chance level [$n = 18$, GLMM: 95% CI = 1.56 (0.72 to 2.40), $Z = 5.90$, $P < 0.001$; **Figure 3** [↗](#); S4 video], and bees spent longer times manipulating connected strings (79.59 ± 10.15 s) [GLMM: 95% CI = 0.09 (0.03 to 0.14), $t = 3.23$, $P < 0.01$; **Figure 4** [↗](#)].

To explore whether bees simply memorized the visual display of the “lollipop shape” of a string connected to a flower during training with a given color combination, we then asked whether bees generalize from the string color used during training. Bumblebees were trained with green strings (S5 video) or blue strings (S6 video) connected to flowers. If the bees had the ability to generalize the function of strings, this color change should not affect their ability to discriminate connected

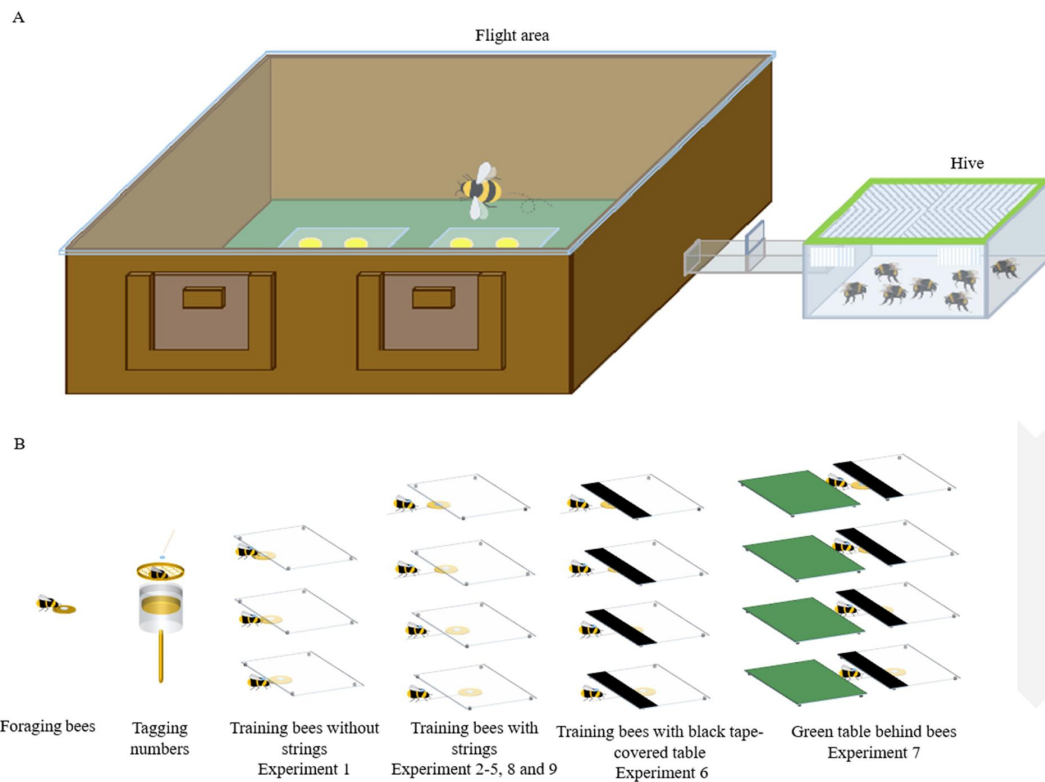


Figure 1.

Experimental apparatus and summary of training in different experiments. (A) The experimental setup consisted of a flight arena connected to a hive via a Perspex corridor. (B) Foraging bumblebees were number-tagged, and the marked bees were trained in a stepwise manner.

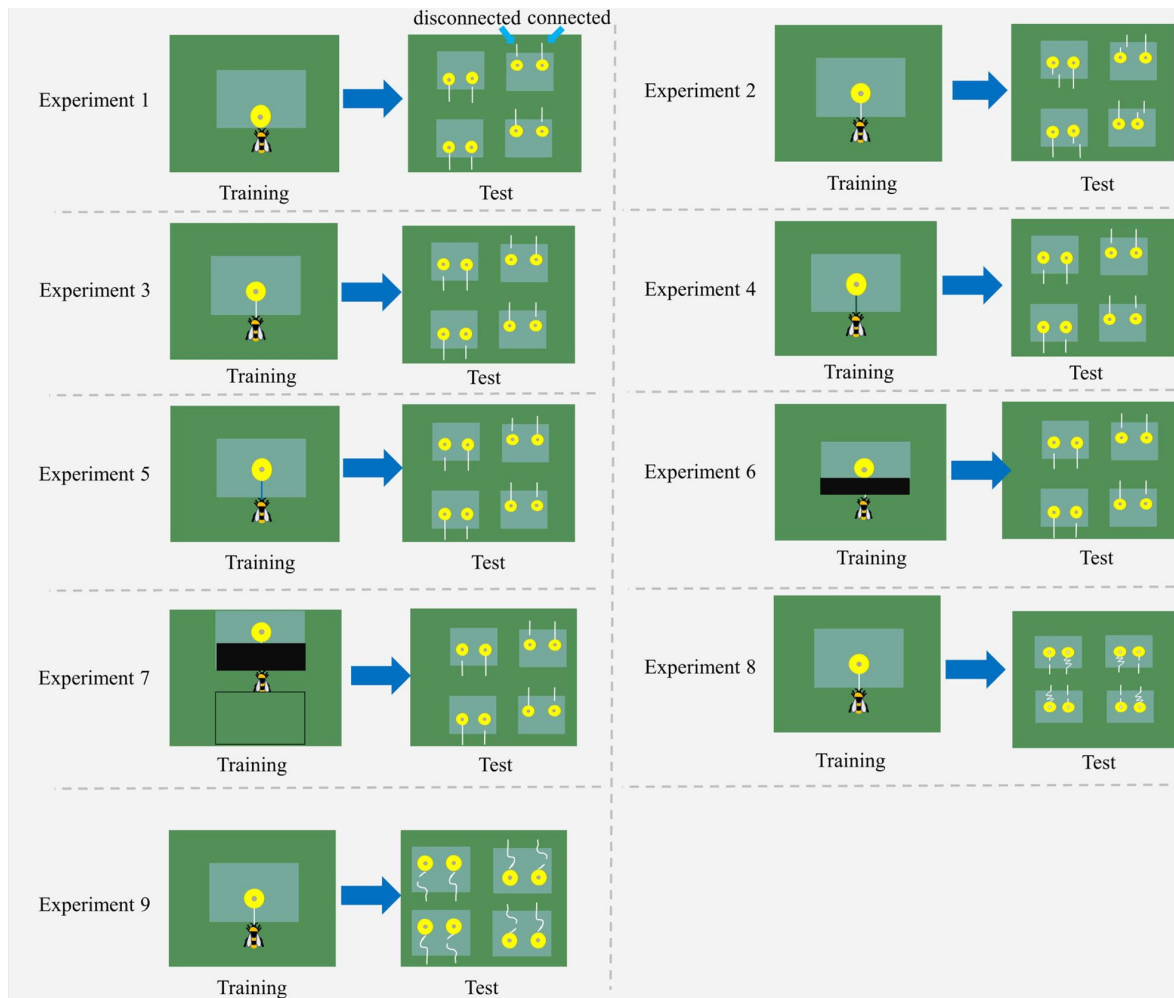


Figure 2.

Schemes of nine string-pulling experiments. Bumblebees were trained and tested in different situations. Two flowers were placed under each transparent table in the test, with one string connected to a flower and another was detached. For further details and descriptions of each experiment, see the text.

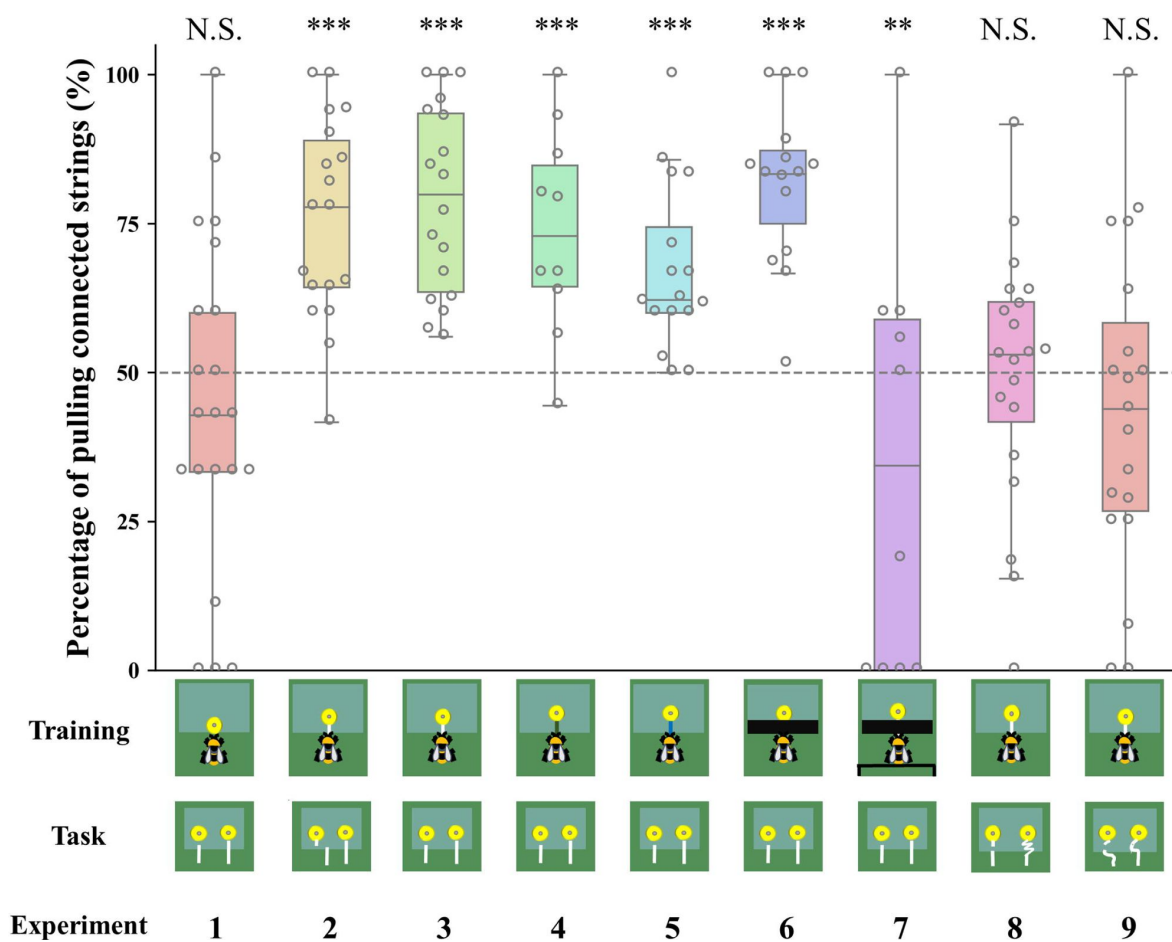


Figure 3.

Bumblebee's preferences for continuous strings over the entire duration of the test in different string-pulling experiments. Percentage of pulling connected strings compared with chance level (50%). Median, interquartile range and range are given. Boxes show the 25th percentile, 50th percentile (median), 75th percentile. The dashed line represents chance level (50%) and the circles indicate individual bees' data points. *** indicates $P < 0.001$; ** - $P < 0.01$; N.S. - $P > 0.05$.

Training									
Task									
Experiment	1	2	3	4	5	6	7	8	9
Number of choices (mean ± SE)	5.10 ± 0.54	15.00 ± 1.53	15.56 ± 2.16	12.10 ± 2.19	11.63 ± 1.66	15.87 ± 2.05	5.60 ± 1.55	20.60 ± 1.80	14.37 ± 2.82
Number of bees choosing connected strings at first choice/N	10/21 (N.S.)	13/18 (N.S.)	17/18 (***)	7/10 (N.S.)	14/16 (*)	10/15 (N.S.)	5/10 (N.S.)	11/20 (N.S.)	10/19 (N.S.)
Chi-square result	$\chi^2 = 0.05$ $P = 0.83$	$\chi^2 = 3.56$ $P = 0.06$	$\chi^2 = 14.22$ $P < 0.001$	$\chi^2 = 1.60$ $P = 0.21$	$\chi^2 = 9.00$ $P < 0.05$	$\chi^2 = 1.67$ $P = 0.20$	$\chi^2 = 0.00$ $P = 1.00$	$\chi^2 = 0.20$ $P = 0.65$	$\chi^2 = 0.05$ $P = 0.82$

Note: The numbers before the slash are the numbers of bumblebees pulling connected strings on the first choice, and the numbers after the slash are the total number of tested bumblebees. *N* is the total number of the bees. *** - $P < 0.001$; * - $P < 0.05$, N.S. - $P > 0.05$.

Table 1.

The number of total choices and bees that chose connected strings at first choice.

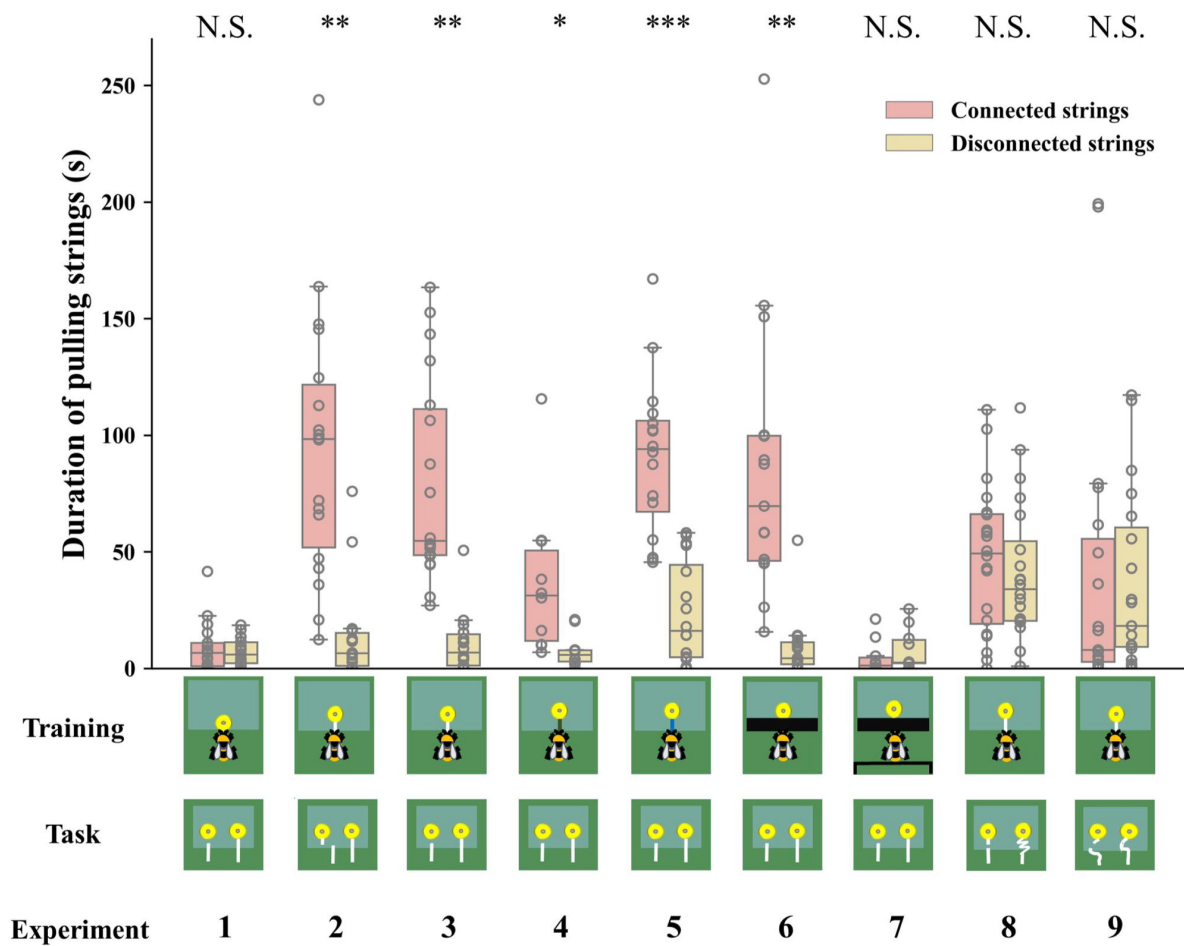


Figure 4.

Duration of bumblebees attempting to pull the strings over the entire duration of the test across experiments. *Median, interquartile range and range are given. Boxes show the 25th percentile, 50th percentile (median), 75th percentile. Circles indicate individual bees' data points. *** - $P < 0.001$; ** - $P < 0.01$; * - $P < 0.05$; N.S. - $P > 0.05$.*

and disconnected strings. Overall, changing the color of the string in training reduced the accuracy of the choice, 7/10 bees (70%, $\chi^2 = 1.60$, $P = 0.21$) selected the connected string in the first choice (**Table 1**). But over the entire test, the bees still maintained the basic discrimination at well above chance level when bees trained with green strings ($61 \pm 5\%$) [$n = 10$, generalized linear model (GLM): 95% CI = 1.13 (0.71 to 1.54), $Z = 5.27$, $P < 0.001$; **Figure 3**]. In addition, the duration of pulling connected strings (47.79 ± 9.91 s) was significantly longer than disconnected ones (24.78 ± 6.05 s) [$n = 10$, GLMM: 95% CI = 0.10 (0.01 to 0.19), $t = 2.26$, $P < 0.05$; **Figure 4**]. In this sense, bees might possess the ability to generalize string color from the originally learned stimulus. A similar result was found in bees trained with blue strings. We found that 14 of 16 bees (87%, $\chi^2 = 9.00$, $P < 0.05$) selected the connected string in the first choice (**Table 1**). The percentage of bees pulling connected strings was significantly higher than chance level [$n=16$, GLMM: 95% CI = 0.60 (0.30 to 0.90), $Z = 3.90$, $P < 0.001$; **Figure 3**], and the duration data also indicate that the bees prefer pulling connected strings [$n=16$, linear mixed model (LMM): 95% CI = -66.91 (-80.86 to -52.96), $t = -9.68$, $P < 0.001$; **Figure 4**].

The question of whether animals rely on perceptual feedback during string pulling has been tested with occluders between the string and the reward in a variety of other species (Taylor *et al.*, 2010; Gaycken *et al.*, 2019; Molina *et al.*, 2019; Wakonig *et al.*, 2021). Bumblebees were trained to feed on yellow artificial flowers, and then trained with transparent tables covered by black tape through a four-step process (S7 video). The aim was to prevent the bees from seeing the string connected to the flower during training. Note, however, the bees were able to see this “lollipop shape” (string connected to the flower) after they pulled the strings out from the table or during the initial step of the training (**Figure 1**). Ten out of fifteen bees (67%, $\chi^2 = 1.67$, $P = 0.20$) pulled the connected string in their first choice (**Table 1**). However, over the full duration of the test, the percentage of the bees pulling connected strings ($82 \pm 3\%$) was significantly higher than chance level [$n=15$, GLMM: 95% CI = 1.31(1.00 to 1.62), $Z = 8.23$, $P < 0.001$; **Figure 3**], and the duration data also indicate that the bees preferred pulling connected strings [$n=15$, GLMM: 95% CI = 0.10 (0.02 to 0.17), $t = 2.61$, $P < 0.01$; **Figure 4**]. To help ensure that the bees could not see the “lollipop shape” during the training phase, we placed a green table behind the bee (**Figure 2**). In this way, the “lollipop shape” was not directly presented during the initial step of training, nor was it visible after the bee had reached the reward because the string was then covered by the back green table (S8 video). We found that the initial choices of bees in this test were at chance level (5 out of 10 bees chose the connected string). Over the full test, the percentage of pulling connected strings was significantly lower than chance level [$n=10$, GLMM: 95% CI = -0.77 (-1.33 to -0.20), $Z = -2.61$, $P < 0.01$; **Figure 3**]. But duration data indicated that the bees showed no preference for pulling the connected strings or disconnected strings [$n=10$, GLMM: 95% CI = -0.09 (-0.38 to 0.19), $t = -0.64$, $P = 0.53$; **Figure 4**].

So far, our results show that bumblebees could have been using image matching to discriminate connected from disconnected strings in the test. We therefore designed further experiments based on Taylor *et al.* (2012) to test this hypothesis. Bumblebees were first trained to feed on yellow artificial flowers, and then trained with the same procedure as Experiment 2, but the connected strings were coiled in the test. Bees failed to choose the connected strings when strings were coiled. We observed 11/20 bees (55%, $\chi^2 = 0.20$, $P = 0.65$) pulled the connected string in their first choice (**Table 1**). Over the full duration of the test, no difference in the percentage of pulling connected strings compared with chance level [$n=20$, GLMM: 95% CI = 0.17 (-0.26 to 0.59), $Z = 0.70$, $P = 0.48$; **Figure 3**]. There was also no significant difference between the duration of bees pulling the connected strings and disconnected strings [$n=20$, LMM: 95% CI = -5.77 (-24.04 to 12.50), $t = -0.63$, $P = 0.53$; **Figure 4**], suggesting that bumblebees did not recognize the continuity of coiled strings. Another group of bees was trained with straight strings, while both connected and disconnected strings were coiled in the test. Ten of nineteen bees (53%, $\chi^2 = 0.05$, $P = 0.82$) chose the connected string as their first choice in the test (**Table 1**). Similarly, no differences

were found in percentage [$n=19$, GLMM: 95% CI = -0.03 (-0.54 to 0.48), $Z = -0.29$, $P = 0.77$; **Figure 3**] and duration [$n=19$, GLMM: 95% CI = 0.005 (-0.02 to 0.03), $t = 0.22$, $P = 0.83$; **Figure 4**] of pulling each string.

Latency to the first choice was defined as the latency to initiate pulling the string after the bee entered the set-up. The latency to the first choice was measured to assess if the bumblebees were familiarizing themselves with the testing pattern. A shorter latency might indicate that the bumblebees were more familiar with the patterns. The latency of the bees that were trained with blue strings (684.33 ± 105.45 s) was substantially longer than that of bees which were trained with white strings and tested with straight strings (**Figure 5**). A long latency time was observed in Experiment 7 (557.26 ± 104.09 s), which the bees were trained with black tape covering table, and green table was placed behind the bees (**Figure 5**).

Discussion

Our results show that: (i) bumblebees require experience with string pulling to distinguish between connected and disconnected strings; (ii) bumblebees are able to generalize features learned during string-pulling training to solve a task with different colored strings; (iii) bumblebees solve string-pulling tasks through image matching.

The results suggest that bumblebees require experience to recognize interrupted strings and acquire a preference for connected ones. This corroborates previous findings that most bees failed to solve single string-pulling tasks without previous training (Alem *et al.*, 2016). Some animals, including dogs (Osthaus *et al.*, 2005), cats (Whitt *et al.*, 2009), western scrub-jays (Hofmann *et al.*, 2016) and azure-winged magpies (Wang *et al.*, 2019) fail in such task spontaneously. It is worth noting that some crows and parrots known for complex cognition perform poorly on the broken string task without perceptual feedback or learning. For example, New Caledonian crows use perceptual feedback strategies to solve the broken string-pulling task, and no individual showed a significant preference for the connected string when perceptual feedback was restricted (Taylor *et al.*, 2012). Some Australian magpies and African grey parrots can solve the broken string task, but they require a high number of trials, indicating that learning plays a crucial role in solving this task (Molina *et al.*, 2019; Johnsson *et al.*, 2023).

Our findings suggest that bumblebees with experience of string pulling prefer the connected strings, but they failed to identify the interrupted strings when the string was coiled in the test. Bees acquire their preferences for flowers with connected strings at least in part by learning the visual appearance of the “lollipop shape” present during training. Trained bees may have memorized this image as a predictor of reward and applied it to solve novel string-pulling tasks. This makes sense because the bees could see the “lollipop shape” once they pulled it out from the table in Experiment 6. Another possibility is that bumblebees might remember the image of the “lollipop shape” from the initial training step, because the shape was directly presented to the bees. However, when a green table was placed behind the string to obscure the “lollipop” structure during the training, the bees could not see the “lollipop” during the initial training stage or after pulling the string from under the table. In this situation, the bees were unable to identify the connected string, further supporting the notion that bumblebees chose the connected string based on image matching. Bumblebees exhibited longer delay times to the first choice in Experiments 5 and 7 (**Figure 5**). The reason might be attributable to bees’ search time for the familiar image or neophobic response (Muller *et al.*, 2010), because the strings during training were changed in test, which the bees had not encountered before.

Bees often need to match memorized images of flowers to currently visible flowers while foraging in the wild (Chittka *et al.* 1999; Giurfa, 2003). Different flower species offer varying profitability in terms of nectar and pollen to bumblebees; they need to make careful choices and

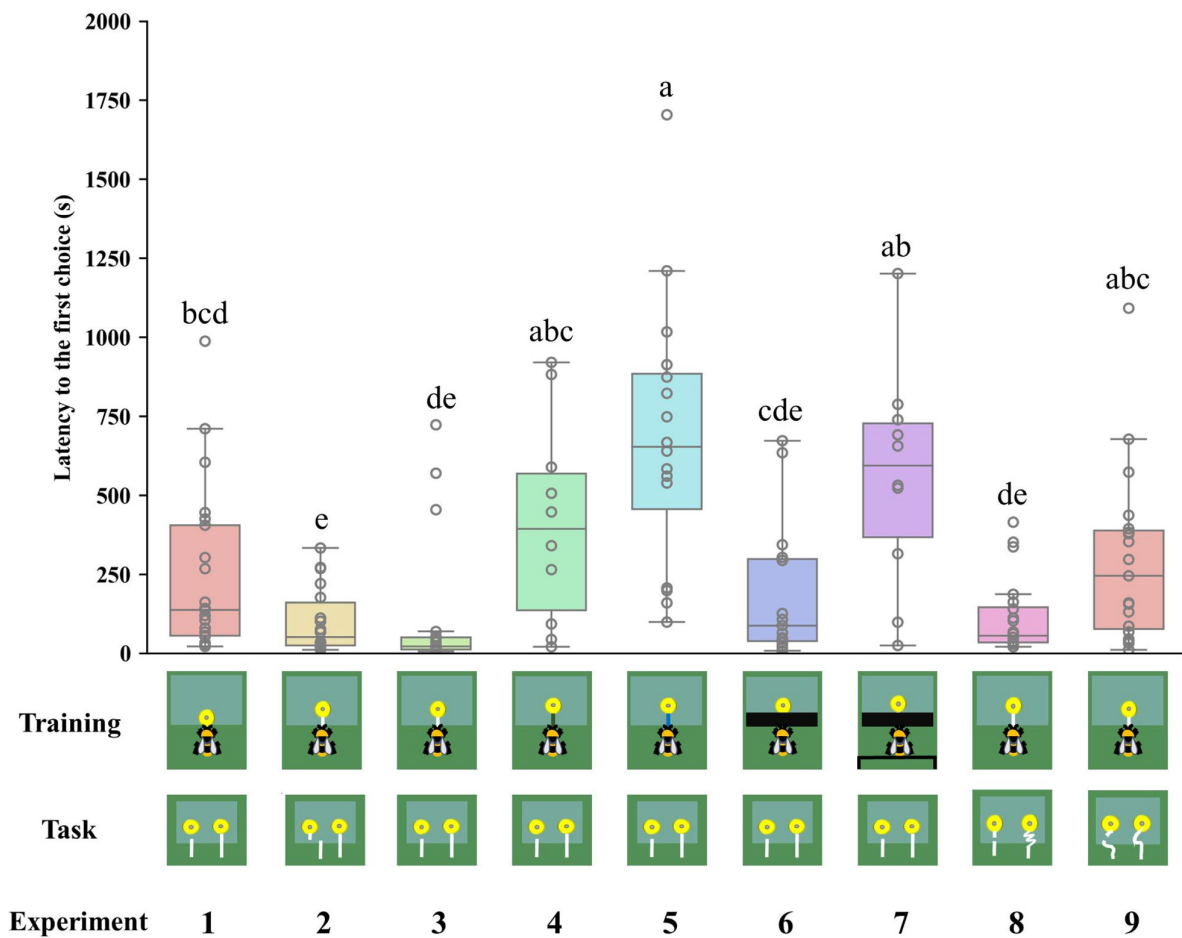


Figure 5.

Latency to the first choice in different experiments. Median, interquartile range and range are given. Boxes show the 25th percentile, 50th percentile (median), 75th percentile. Circles indicate individual bees' data points. Different letters indicate significant differences ($P < 0.05$).

learn to use floral cues to predict rewards (Chittka, 2017). Bumblebees can easily learn visual patterns and shapes of flowers (Meyer-Rochow, 2019); they can detect stimuli and discriminate between differently coloured stimuli when presented as briefly as 25 ms (Nityananda *et al.*, 2014). In contrast, causal reasoning involves understanding and responding to causal relationships. Bumblebees might favor, or be limited to, a visual approach, likely due to the efficiency and simplicity of processing visual cues to solve the string-pulling task.

Rather than relying on means-end comprehension, most animals likely use simpler associative strategies to solve string-pulling tasks, as observed in the bumblebees in the present study (Jacobs & Osvath, 2015; Wakonig *et al.*, 2021). Empirical evidence from several vertebrates has shown that success in object use does not necessarily imply causal understanding; rather, it involves abstracting simple rules based on observable features of the physical task at hand (Seed *et al.*, 2006; Schuck-Paim *et al.*, 2009; Herrmann *et al.*, 2008; Gagne *et al.*, 2012). In several vertebrate string-pulling studies, animals relied on a “proximity rule” in most cases, choosing to pull the strings closest to the reward, regardless of their connectivity (Whitt *et al.*, 2009; Wang *et al.*, 2021).

In conclusion, even though bumblebees may not initially understand the causality of string-pulling, they can match the image of the strings connected to flowers, and rely on associative mechanisms to remember the previously visited stimuli. Bumblebees, whether with or without string-pulling experience, do not appear to understand the value of strings to target objects. This negative result does not necessarily mean that bumblebees are entirely unable to comprehend the link between a means and an end, but our results suggest that for the paradigms tested here, such comprehension is not required.

Materials and methods

Animals and experimental arena

Nineteen colonies of bumblebees (*Bombus terrestris*) each containing a queen were obtained from commercially available stocks provided by a distributor in the United Kingdom (Biobest, Belgium N.V.) or China (Biobest, Belgium N.V.; Biobest Shouguang Biotechnology Co., Ltd). Bees were housed in a plastic nest box (29 × 22.5 × 13.3 cm [L × W × H]) that was connected to a flight arena (100 × 75 × 30 cm [L × W × H]) by an acrylic corridor (25 × 3.5 × 3.5 cm [L × W × H]). The flight arena was covered with an acrylic lid. Three sliding doors were placed along the corridor, allowing the experimenter to control bees' access to the arena (Figure 1A). The floor of the flight arena was painted green, which provided a smooth surface and high-contrast pattern visual panorama between the strings and background for the bees (Spaethe *et al.*, 2001). Outside of experiments, the colonies were provided with 20% (w/w; weight-to-weight) sucrose solution from a gravity feeder placed in the center of the arena, and with ~5g commercially obtained pollen (Koppert B.V., The Netherlands; Changge Yafei Beekeeping Professional Cooperative, China) every other day. All the training programs and tests were conducted in the flight arena between 9 am and 7 pm under light (12 : 12 h) at room temperature (23 ± 4L). Illumination was provided by fluorescent lighting (Osram Sylvania, Wilmington, NC, U.S.A.; YZ36RR, 36W, T8/765, FSL, China) fitted with high-frequency ballasts (HFB 236 TLD, Philips, Amsterdam, The Netherlands; T8, YZ-36 W, FSL, China) to generate lighting above the bee flicker fusion frequency (Skorupski & Chittka, 2010).

Before each experiment, bees were first pretrained to find sucrose solution (50%, w/w) in yellow artificial flowers (3 cm diameter yellow discs with an inverted Eppendorf cap at the center; henceforth “flowers”), which were randomly located in the arena with sucrose solution (50%, w/w) in the Eppendorf cap. Bees that seemed to forage with regularity were number-tagged for individual identification (Figure 1B). In detail, one forager bee was transferred to a cylindrical

cage (diameter=3.8 cm, length=7.7 cm) with a sponge plunger, and a numbered tag (Bienen-Voigt & Warnholz GmbH & Co. KG, Germany) was glued to bee's thorax for individual recognition (**Figure 1B** [↗](#)).

General Methods

For each experiment, bumblebees were trained to retrieve a flower with an inverted Eppendorf cap at the center, containing 25 microliters of 50% sucrose solution, from underneath a transparent acrylic table [0.6 cm above the ground, 15 × 10 × 0.4 cm (L × W × H); henceforth “table”]. For Experiment 1, bees were trained in a stepwise manner – Step 1, 50% of the flower was covered by the transparent acrylic table, Step 2, 75% of the flower was covered; Step 3, 100% of the flower was covered (**Figure 1B** [↗](#)). For Experiment 2-8, the first three steps are similar to Experiment 1, but the flowers were connected to strings (length = 4.5 cm, diameter = 0.3 cm), and bees were trained with the fourth step: 2 cm strings were attached to the flower and accessible from outside the table. The bees received rewards five times in each of the steps, except for the last step. The training phase was completed when a bee pulled the strings and drank from the flowers twenty times after the first occurrence of string pulling during the last step. For each test, bees were individually tested in the arena and presented with four transparent tables. Two options were placed under each table, parallel to each other, and perpendicular to the long side of the table. To avoid developing a side bias, the position and direction of the strings varied randomly from left to right for each table. During tests, both strings were glued to the floor of the arena to prevent the air flow generated by flying bumblebees' wings from changing the position of the string. Different groups of bees were tested for Experiments 1-9, and each forager bee was used only once, and the tested bees were removed from the nest and then placed in the freezer to be euthanized. All the experiments were videotaped with an iPhone 12 (Apple, Cupertino, CA, USA) placed above the arena. The choice and duration of the bees pulling the connected or disconnected strings were recorded. A choice was recorded when a bee used her legs or mandibles to pull the connected or disconnected strings. The test was terminated when the bee stopped engaging with the tables, flowers and strings for more than one minute, and a testing session lasted a maximum of 30 min.

Experiment 1: Do bumblebees without string-pulling experience discriminate between connected and disconnected strings?

Bumblebees (n=21) were trained with artificial flowers under a transparent table. Initially, half of the flower was placed under the table, and the central Eppendorf cap (containing the sucrose solution reward) at the edge of the table so that the bees could access sugar water directly without moving the flower (**Figure 1B** [↗](#)). In the final step, the edge of the flower was aligned with the table, and bees had to pull the edge of the flower to obtain the reward (**Figure 1B** [↗](#)). In the test, four transparent tables were placed on the arena floor, and two artificial flowers were placed 3 cm apart under each table, with one connected to a string (length=2.5 cm, including 1 cm accessible from outside the table, diameter=0.3 cm), while another flower was presented with a 1.5 cm string with a cm gap between the string segment and the flower (**Figure 2** [↗](#)). If the bees prefer to pull the connected strings, this would indicate that bees naturally recognize the connectivity of strings in this task.

Experiments 2-3: Do bumblebees with string-pulling experience discriminate between connected and disconnected strings?

In Experiment 2, bumblebees (n=18) were trained to pull a white string (length = 4.5 cm, diameter = 0.3 cm) attached to the yellow artificial flower that was placed under a transparent table, following the stepwise string-pulling training protocol used by [Alem et al. \(2016\)](#) [↗](#) with some modifications. Briefly, selected bees were trained to pull the string when the flowers were gradually positioned further under the table, finally, the strings protruded 2 cm outside the table

edge (**Figure 1B** [↗](#)). In the test, four tables were placed on the arena floor, two artificial flowers were placed 3 cm apart under each table, with one flower connected to a long string (length=4.5 cm), whereas another flower was attached to a short string (length=1 cm), and a 3.5 cm string segment (2 cm accessible from outside the table) was positioned with a lateral displacement. The short string and string segment were placed along parallel lines and 1.5 cm apart from each other (**Figure 2** [↗](#)). The position of the string segment was randomly assigned on the left or right side of the short string.

In Experiment 3, a further 18 bumblebees were trained with the same procedure as Experiment 2. In the test, four tables were placed on the arena floor, two artificial flowers were placed 3 cm apart under each table, with one connected to a long string (length=4.5 cm), a 3.5 cm string segment was placed along with another flower, and a 1 cm gap was created between the string segment and flower (**Figure 2** [↗](#)). If the bees show a preference for the strings connected to flowers, this would indicate that bees with string-pulling experience can recognize connected and disconnected strings.

Experiments 4-5: Do bumblebees generalize when string colors differ between training and testing?

In Experiment 4, to further verify whether bees can generalize from the string color used during training, bees (n=21) were trained with yellow flowers connected to green strings (length=4.5 cm), such that strings were visually different from the white strings in the test (**Figure 2** [↗](#)). The training method was the same as Experiment 2 and the test protocol was the same as for Experiment 3. If the bees prefer to pull the connected strings, indicating that bees generalize the color of the strings in the training. In Experiment 5, bumblebees (n=16) were trained with yellow flowers connected to blue strings (length=4.5 cm) (**Figure 2** [↗](#)). The training method was the same as for Experiment 2. The test protocol was the same as for Experiment 3. If the bees prefer to pull the connected strings, this indicated the bees can generalize from the string color used during training.

Experiment 6-9: Do bumblebees use image matching to discriminate connected versus disconnected strings?

In Experiment 6, bumblebees (n=15) were trained with a transparent Perspex table where the edge on which the strings protruded was covered with opaque black tape (15 × 2.5 cm [L × W]) (**Figure 1B** [↗](#)). The bees could not see the strings above the table; however, they could receive visual feedback both after the string was pulled out from the table and during the initial stages of training. The test protocol was the same as for Experiment 3. If the bees prefer to pull the connected strings, this would indicate that bees memorize the arrangement of strings-connected flowers in this task.

In Experiment 7, bumblebees (n=22) were trained with a piece of black tape covering the front part of the transparent table, and a table completely covered with a green board was placed behind the bees (**Figure 1B** [↗](#)). The distance between the two tables was 2 cm. The entire string was not visible during the initial step of training, and the string disappeared from the bees' view after being pulled out from the table. The test protocol was the same as in Experiment 3. If the bees failed the connectivity task, it would indicate that they used visual cues to solve this task. We trained 22 bees, but 12 of them did not pull the string even once during the test. The bees that did not pull the string were not included in the statistical analysis.

In Experiment 8, bumblebees (n=20) were trained with the same protocol as for Experiment 2. In the test, two artificial flowers were placed 3 cm apart under each table, with one flower connected to a coiled string (length=20 cm, four turns), this pattern was visually different from a straight string during training. Another flower was attached to a short string (length=1 cm), a 2.5 cm

straight string segment was placed along the short string (**Figure 2**), and a 1 cm gap between the short string and the 3.5 cm string segment was presented. If the bees show a preference for the straight strings (same arrangement in training) over the coiled string or show no preference for either of the two strings, this would indicate that they use image matching to solve this task.

In Experiment 9, bumblebees ($n=19$) were trained with the same protocol as for Experiment 2. In the test, two artificial flowers were placed under each table, with one flower connected to a coiled string (length=8 cm, one turn). The other flower was attached to a short string (1 cm), a coiled string segment (length=6 cm, one turn) was placed along the short string, and there was a 1 cm gap between the flower and the 6.5 cm coiled string segment (**Figure 2**). The two coiled strings in the test were visually different from a straight string in training. If bumblebees use image matching to discriminate connected and disconnected strings, they will exhibit no preference for either of the two strings.

Statistical Analyses

All statistical analyses were conducted with R version 4.2.0. In each experiment, the percentage of pulling connected strings of the bees was analyzed with generalized linear mixed models (GLMM) [R Development Core Team, lme4 package]. Colony (>1) was set as random effect. If the bees in an experiment were from a single colony (Experiment 4), the percentage was analyzed with a generalized linear model (GLM). Binomial distribution and logit function were employed for both models. The total number of choices made by each bee was set as weights (Bates *et al.*, 2015). Bees' first choices between connected and disconnected strings were analyzed with Chi-square tests. The duration of pulling different kinds of strings and the switching interval were first tested with the Shapiro-Wilk test to assess data normality. The duration data that conform to a normal distribution were compared using linear mixed-effects models (LMM), while the data that deviated from normality were examined using a generalized linear-mixed model (GLMM). In each model, the duration were set as the dependent variable, and string type was considered as a fixed effect, and the bee identity and colony (>1) as random effects. Latency to the first choice was analyzed with GLMM, where the experiment was set as a fixed effect and the colony as a random effect with gamma family. The emmeans package (Lenth *et al.*, 2024) was employed to conduct multiple comparisons among different experiments.

Acknowledgements

We thank Jian Chen (Usda-Ars, Biological Control of Pests Research Unit, Stoneville, Mississippi) for valuable comments and suggestions on the early version of the manuscript. This work was supported by the National Natural Science Foundation of China (32301292) and National Natural Science Foundation of China (32271888). Cwyn Solvi was supported by a Templeton World Charity Foundation project grant (TWCF-2020-0539).

Authors' contributions

Chao Wen: Conceptualization, Resources, Data curation, Formal analysis, Supervision, Funding acquisition, Investigation, Writing-original draft, Project administration, Writing-review and editing; Yuyi Lu: Resources, Data curation, Formal analysis, Investigation, Methodology, Writing-review and editing; Cwyn Solvi, Writing-review and editing, Methodology; Shunping Dong: Investigation, Formal analysis, Methodology. Cai Wang, Xiujun Wen, Haijun Xiao Shikui Dong, Junbao Wen: Supervision, Resources, Writing-review and editing; Fei Peng, Lars Chittka: Conceptualisation, Supervision, Methodology, Funding acquisition, Validation, Visualization, Project administration, Writing-original draft, Writing-review and editing.

Competing interests

We declare we have no competing interests.

References

- Alem S, Perry CJ, Zhu XF, Loukola OJ, Ingraham T, Søvik E, Chittka L (2016) **Associative mechanisms allow for social learning and cultural transmission of string pulling in an insect** *PLoS Biology* **14** <https://doi.org/10.1371/journal.pbio.1002589>
- Bates D, Mächler M, Bolker B, Walker S (2015) **Fitting Linear Mixed-Effects Models Using lme4** *Journal of Statistical Software* **67**:1–48 <https://doi.org/10.18637/jss.v067.i01>
- Chittka L. (2017) **Bee cognition** *Current Biology* **27**:1049–1053
- Chittka L, Rossiter SJ, Skorupski P, Fernando C (2012) **What is comparable in comparative cognition?** *Philosophical Transactions Of The Royal Society B-biological Sciences* **367**:2677–2685 <https://doi.org/10.1098/rstb.2012.0215>
- Chittka L, Thomson JD, Waser NM (1999) **Flower constancy, insect psychology, and plant evolution** *Naturwissenschaften* **86**:361–377
- de Waal FBM, Ferrari PF (2010) **Towards a bottom-up perspective on animal and human cognition** *Trends in Cognition Science* **14**:201–207 <https://doi.org/10.1016/j.tics.2010.03.003>
- Gagne M, Levesque K, Nutile L, Locurto C (2012) **Performance on patterned string problems by common marmosets (*Callithrix jacchus*)** *Animal Cognition* **15**:1021–1030 <https://doi.org/10.1007/s10071-012-0511-0>
- Gaycken J, Picken DJ, Pike TW, Burman OHP, Wilkinson A (2019) **Mechanisms underlying string-pulling behaviour in green-winged macaws** *Behaviour* **156**:619–631 <https://doi.org/10.1163/1568539X-00003520>
- Giurfa M (2003) **Cognitive neuroethology: dissecting non-elemental learning in a honeybee brain** *Current Opin in Neurobiology* **13**:726–735 <https://doi.org/10.1016/j.conb.2003.10.015>
- Herrmann E, Wobber V, Call J (2008) **Great apes' (*Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, *Pongo pygmaeus*) understanding of tool functional properties after limited experience** *Journal of Comparative Psychology* **122**:220–230 <https://doi.org/10.1037/0735-7036.122.2.220>
- Hofmann MM, Cheke LG, Clayton NS (2016) **Western scrub-jays (*Aphelocoma californica*) solve multiple-string problems by the spatial relation of string and reward** *Animal Cognition* **19**:1103–1114 <https://doi.org/10.1007/s10071-016-1018-x>
- Jacobs IF, Osvath M (2015) **The string-pulling paradigm in comparative psychology** *Journal of Comparative Psychology* **129**:89–120 <https://doi.org/10.1037/a0038746>
- Johnsson RD, Veillet PS, Connelly F, Endler JA, Roth TC, Lesku JA (2023) **Wild Australian magpies learn to pull intact, not broken, strings to obtain food** *Behavioral Ecology and Sociobiology* **77** <https://doi.org/10.1007/s00265-023-03326-6>

- Lenth RV, Bolker B, Buerkner P, Giné-Vázquez I, Herve M, Jung M, Love J, Miguez F., Riebl H, Singmann H (2024) **Lenth RV, Bolker B, Buerkner P, Giné-Vázquez I, Herve M, Jung M, Love J, Miguez F., Riebl H, Singmann H. 2024. emmeans: Estimated Marginal Means, aka Least-Squares Means (1.10.0) [Computer software].** (<https://cran.r-project.org/web/packages/emmeans/index.html>)
- Mayer C, Call J, Aliach-Serrano A, Visalerghi E, Sabbatini G, Seed A (2014) **Abstract knowledge in the broken-string problem: evidence from nonhuman primates and pre-schoolers** *PLoS One* **9** <https://doi.org/10.1371/journal.pone.0108597>
- Molina ABC, Cullell TM, Mimó MC (2019) **String-pulling in African grey parrots (*Psittacus erithacus*): Performance in discrimination tasks** *Behaviour* **156**:847–857 <https://doi.org/10.1163/1568539X-00003511>
- Meyer-Rochow VB (2019) **Eyes and vision of the bumblebee: a brief review on how bumblebees detect and perceive flowers** *Journal of Apiculture* **34**:107–155 <https://doi.org/10.1242/jeb.101394>
- Muller H, Grossmann H, Chittka L (2010) **‘Personality’ in bumblebees: individual consistency in responses to novel colours?** *Animal Behaviour* **80**:1065–1074 <https://doi.org/10.1016/j.anbehav.2010.09.016>
- Nityananda V, Skorupski P, Chittka (2014) **Can bees see at a glance?** *Journal of Experimental Biology* **217**:1933–1939 <https://doi.org/10.1242/jeb.101394>
- Ortiz ST, Maxwell A, Krasheninnikova A, Wahlberg M, Larsen ON (2019) **Problem solving capabilities of peach-fronted conures (*Eupsittula aurea*) studied with the string-pulling test** *Behaviour* **156**:815–846 <https://doi.org/10.1163/1568539X-00003539>
- Osthaus B, Lea SE, Slater AM (2005) **Dogs (*Canis lupus familiaris*) fail to show understanding of means-end connections in a string-pulling task** *Animal Cognition* **8**:37–47 <https://doi.org/10.1007/s10071-004-0230-2>
- Piaget J (1953) **The origins of intelligence in children** <https://doi.org/10.1037/11494-000>
- Range F, Möslinger H, Virányi Z (2012) **Domestication has not affected the understanding of means-end connections in dogs** *Animal Cognition* **15**:597–607 <https://doi.org/10.1007/s10071-012-0488-8>
- Riemer S, Müller C, Range F, Huber L (2013) **Dogs (*Canis familiaris*) can learn to attend to connectivity in string pulling tasks** *Journal of Comparative Psychology* **128**:31–39 <https://doi.org/10.1037/a0033202>
- Schuck-Paim C, Borsari A, Ottoni EB (2009) **Means to an end: neotropical parrots manage to pull strings to meet their goals** *Animal Cognition* **12**:287–301 <https://doi.org/10.1007/s10071-008-0190-z>
- Seed AM, Tebbich S, Emery NJ, Clayton NS (2006) **Investigating physical cognition in rooks, *Corvus frugilegus*** *Current Biology* **16**:697–701 <https://doi.org/10.1016/j.cub.2006.02.066>
- Shettleworth SJ (2010) **Clever animals and killjoy explanations in comparative psychology** *Trends in Cognition Science* **14**:477–481 <https://doi.org/10.1016/j.tics.2010.07.002>

Skorupski P, Chittka L (2010) **Differences in photoreceptor processing speed for chromatic and achromatic vision in the bumblebee, *Bombus terrestris*** *The Journal of Neuroscience* **30**:3896–3903 <https://doi.org/10.1523/JNEUROSCI.5700-09.2010>

Spaethe J, Tautz J, Chittka L (2001) **Visual constraints in foraging bumblebees: flower size and color affect search time and flight behavior** *Proceedings of the National Academy of Sciences of the United States of America* **98**:3898–903

Taylor AH, Knaebe B, Gray RD (2012) **An end to insight? New Caledonian crows can spontaneously solve problems without planning their actions** *Proceedings of the Royal Society B: Biological Sciences* **279**:4977–4981

Taylor AH, Medina FS, Holzhaider JC, Hearne LJ, Hunt GR, Gray RD (2010) **An investigation into the cognition behind spontaneous string pulling in new caledonian crows** *PLoS One* **5** <https://doi.org/10.1371/journal.pone.0009345>

Wakonig B, Auersperg AMI, O'Hara M (2021) **String-pulling in the Goffin's cockatoo (*Cacatua goffiniana*)** *Learning & Behavior* **49**:124–136 <https://doi.org/10.3758/s13420-020-00454-1>

Wang L, Luo YC, Wang X, Maierdiyali A, Chang H, Li ZQ (2019) **Azure-winged magpies solve string-pulling tasks by partial understanding of the physical cognition** *Current Zoology* **65**:385–392 <https://doi.org/10.1093/cz/zoy070>

Wang L, Zhang DJ, Sui JL (2021) **Investigation of cognitive mechanisms and strategy on solving multiple string-pulling problems in Azure-winged magpie (*Cyanopica cyanus*)** *Animal Cognition* **4**:1–10 <https://doi.org/10.1007/s10071-020-01413-z>

Wasserman EA, Nagasaka Y, Castro L, Brzykcy SJ (2013) **Pigeons learn virtual patterned-string problems in a computerized touch screen environment** *Animal Cognition* **16**:737–753 <https://doi.org/10.1007/s10071-013-0608-0>

Whitt E, Douglas M, Osthaus B, Hocking L (2009) **Domestic cats (*Felis catus*) do not show causal understanding in a string-pulling task** *Animal Cognition* **12**:739–743 <https://doi.org/10.1007/s10071-009-0228-x>

Zhou DB, Dong SP, Ge J, Chittka L, Wang C, Wen C, Wen JB (2024) **Bumblebees attend to both the properties of the string and the target in string-pulling tasks, but prioritize the features of the string** *Insect Science* <https://doi.org/10.1111/1744-7917.13373>

Editors

Reviewing Editor

Rosalyn Gloag

University of Sydney, Sydney, Australia

Senior Editor

Detlef Weigel

Max Planck Institute for Biology Tübingen, Tübingen, Germany

Reviewer #1 (Public Review):

Summary:

In this paper the researchers aimed to address whether bees causally understand string-pulling through a series of experiments. I first briefly summarize what they did:

- In experiment 1, the researchers trained bees without string and then presented them with flowers in the test phase that either had connected or disconnected strings, to determine what their preference was without any training. Bees did not show any preference.
- In experiment 2, bees were trained to have experience with string and then tested on their choice between connected vs. disconnected string.
- Experiment 3 was similar except that instead of having one option which was an attached string broken in the middle, the string was completely disconnected from the flower.
- In experiment 4, bees were trained on green strings and tested on white strings to determine if they generalize across color.
- In experiment 5, bees were trained on blue strings and tested on white strings.
- In experiment 6, bees were trained where black tape covered the area between the string and the flower (i.e. so they would not be able to see/ learn whether it was connected or disconnected).
- In experiments 2-6, bees chose the connected string in the test phase.
- In experiment 7, bees were trained as in expt 3 and then tested where string was either disconnected or coiled i.e. still being 'functional' but appearing different.
- In experiment 8, bees were trained as before and then tested on string that was in a different coiled orientation, either connected or disconnected.
- In experiments 7 and 8 the bees showed no preference.

Strengths:

I appreciate the amount of work that has gone into these experiments and think they are a nice, thorough set of experiments. I enjoyed reading the paper and felt that it was overall well-written and clear. I think experiment 1 shows that bees do not have an untrained understanding of the function of the string in this context. The rest of the experiments indicate that with training, bees have a preference for unbroken over broken string and likely use visual cues learned during training to make this choice. They also show that as in other contexts, bees readily generalize across different colors.

The 'weaknesses' that I previously listed were dealt with by the authors in the revised version of the manuscript. I think the only point that we disagreed on was relating to the ecological relevance of the task to the bees.

Here is my previous comment:

I think the paper would be made stronger by considering the natural context in which the bee performs this behavior. Bees manipulate flowers in all kinds of contexts, and scrabble with their legs to achieve nectar rewards. Rather than thinking that it is pulling a string, my guess would be that the bee learns that a particular motor pattern within their usual foraging repertoire (scrabbling with legs), leads to a reward. I don't think this makes the behavior any

less interesting - in fact, I think considering the behavior through an ecological lens can help make better sense of it.

The authors disagreed, writing the following:

"Here we respectfully disagree. The solving of Rubik's cube by humans could be said to be a version of finger movements naturally required to open nuts or remove ticks from fur, but this is somewhat beside the point: it's not the motor sequences that are of interest, but the cognition involved. A general approach in work on animal intelligence and cognition is to deliberately choose paradigms that are outside the animals' daily routines; this is what we have done here, in asking whether there is means-end comprehension in bee problem solving. Like comparable studies on this question in other animals, the experiments are designed to probe this question, not one of ecological validity."

I think the difference would be that humans know that they are doing a Rubik's cube whereas I do not think that the bee knows that it is pulling string - I think the bee thinks that it is foraging on a flower. Therefore, I stand by my statement that I think it's worth considering what the bee is experiencing in this task and how it relates to what it would be doing while foraging. I think that as animal cognition researchers we can design tasks that are distinct from what the animal would naturally encounter to ask specific questions about what they are thinking - but that we can never remove the ecological context since the animal will always be viewing the task through that lens. However, I think this may be a philosophical difference in opinion and I am happy with the manuscript as it stands.

<https://doi.org/10.7554/eLife.97018.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1 (Public Review):

Summary:

In this paper, the researchers aimed to address whether bees causally understand string-pulling through a series of experiments. I first briefly summarize what they did:

- In experiment 1, the researchers trained bees without string and then presented them with flowers in the test phase that either had connected or disconnected strings, to determine what their preference was without any training. Bees did not show any preference.*
- In experiment 2, bees were trained to have experience with string and then tested on their choice between connected vs. disconnected string.*
- Experiment 3 was similar except that instead of having one option which was an attached string broken in the middle, the string was completely disconnected from the flower.*
- In experiment 4, bees were trained on green strings and tested on white strings to determine if they generalize across color.*
- In experiment 5, bees were trained on blue strings and tested on white strings.*

- In experiment 6, bees were trained where black tape covered the area between the string and the flower (i.e. so they would not be able to see/ learn whether it was connected or disconnected).
- In experiments 2-6, bees chose the connected string in the test phase.
- In experiment 7, bees were trained as in experiment 3 and then tested where the string was either disconnected or coiled i.e. still being 'functional' but appearing different.
- In experiment 8, bees were trained as before and then tested on a string that was in a different coiled orientation, either connected or disconnected.
- In experiments 7 and 8 the bees showed no preference.

Strengths:

I appreciate the amount of work that has gone into this study and think it contains a nice, thorough set of experiments. I enjoyed reading the paper and felt that overall it was well-written and clear. I think experiment 1 shows that bees do not have an untrained understanding of the function of the string in this context. The rest of the experiments indicate that with training, bees have a preference for unbroken over broken string and likely use visual cues learned during training to make this choice. They also show that as in other contexts, bees readily generalize across different colors.

Weaknesses:

(1) I think there are 2 key pieces of information that can be taken from the test phase - the bees' first choice and then their behavior across the whole test. I think the first choice is critical in terms of what the bee has learned from the training phase - then their behavior from this point is informed by the feedback they obtain during the test phase. I think both pieces of information are worth considering, but their behavior across the entire test phase is giving different information than their first choice, and this distinction could be made more explicit. In addition, while the bees' first choice is reported, no statistics are presented for their preferences.

We agree with the reviewer that the first choice is critical in terms of what the bumblebees have learned from the training phase. We analyzed the bees' first choice in Table 1, and we added the tested videos. The entire connected and disconnected strings were glued to the floor, the bees were unable to move either the connected or disconnected strings, and avoid learning behavior during the tests. We added the data of bee's each choice in the Supplementary table.

(2) It seemed to me that the bees might not only be using visual feedback but also motor feedback. This would not explain their behavior in the first test choice, but could explain some of their subsequent behavior. For example, bees might learn during training that there is some friction/weight associated with pulling the string, but in cases where the string is separated from the flower, this would presumably feel different to the bee in terms of the physical feedback it is receiving. I'd be interested to see some of these test videos (perhaps these could be shared as supplementary material, in addition to the training videos already uploaded), to see what the bees' behavior looks like after they attempt to pull a disconnected string.

We added supplementary videos of testing phase. As noted in General Methods, both connected and disconnected strings were glued to the floor to prevent the air flow generated by flying bumblebees' wings from changing the position of the string during the testing phase. The bees were unable to move either the connected or disconnected strings during the

tests, and only attempted to pull them. Therefore, the difference in the friction/weight of pulling the both strings cannot be a factor in the test.

(3) *I think the statistics section needs to be made clearer (more in private comments).*

We changed the statistical analysis section as suggested by the reviewer.

(4) *I think the paper would be made stronger by considering the natural context in which the bee performs this behavior. Bees manipulate flowers in all kinds of contexts and scrabble with their legs to achieve nectar rewards. Rather than thinking that it is pulling a string, my guess would be that the bee learns that a particular motor pattern within their usual foraging repertoire (scrabbling with legs), leads to a reward. I don't think this makes the behavior any less interesting - in fact, I think considering the behavior through an ecological lens can help make better sense of it.*

Here we respectfully disagree. The solving of Rubik's cube by humans could be said to be version of finger-movements naturally required to open nuts or remove ticks from fur, but this is somewhat beside the point: it's not the motor sequences that are of interest, but the cognition involved. A general approach in work on animal intelligence and cognition is to deliberately choose paradigms that are outside the animals' daily routines-this is what we have done here, in asking whether there is means-end comprehension in bee problem solving. Like comparable studies on this question in other animals, the experiments are designed to probe this question, not one of ecological validity.

Reviewer #2 (Public Review):

Summary:

The authors wanted to see if bumblebees could succeed in the string-pulling paradigm with broken strings. They found that bumblebees can learn to pull strings and that they have a preference to pull on intact strings vs broken ones. The authors conclude that bumblebees use image matching to complete the string-pulling task.

Strengths:

The study has an excellent experimental design and contributes to our understanding of what information bumblebees use to solve a string-pulling task.

Weaknesses:

Overall, I think the manuscript is good, but it is missing some context. Why do bumblebees rely on image matching rather than causal reasoning? Could it have something to do with their ecology? And how is the task relevant for bumblebees in the wild? Does the test translate to any real-life situations? Is pulling a natural behaviour that bees do? Does image matching have adaptive significance?

We appreciate the valuable comment from the reviewer. Our explanation, which we have now added to the manuscript, is as follows:

“Different flower species offer varying profitability in terms of nectar and pollen to bumblebees; they need to make careful choices and learn to use floral cues to predict rewards (Chittka, 2017). Bumblebees can easily learn visual patterns and shapes of flower (Meyer-Rochow, 2019); they can detect stimuli and discriminate between differently coloured stimuli when presented as briefly as 25 ms (Nityananda et al., 2014). In contrast, causal reasoning involves understanding and responding to causal relationships. Bumblebees might

favor, or be limited to, a visual approach, likely due to the efficiency and simplicity of processing visual cues to solve the string-pulling task.”

As above, it worth noting that our work is not designed as an ecological study, but one about the question of whether causal reasoning can explain how bees solve a string-pulling puzzle. We have a cognitive focus, in line with comparable studies on other animals. We deliberately chose a paradigm that is to some extent outside of the daily challenges of the animal.

Reviewer #3 (Public Review):

Summary:

This paper presents bees with varying levels of experience with a choice task where bees have to choose to pull either a connected or unconnected string, each attached to a yellow flower containing sugar water. Bees without experience of string pulling did not choose the connected string above chance (experiment 1), but with experience of horizontal string pulling (as in the right-hand panel of Figure 4) bees did choose the connected string above chance (experiments 2-3), even when the string colour changed between training and test (experiments 4-5). Bees that were not provided with perceptual-motor feedback (i.e they could not observe that each pull of the string moved the flower) during training still learned to string pull and then chose the connected string option above chance (experiment 6). Bees with normal experience of string pulling then failed to discriminate between connected and unconnected strings when the strings were coiled or looped, rather than presented straight (experiments 7-8).

Weaknesses:

The authors have only provided video of some of the conditions where the bees succeeded. In general, I think a video explaining each condition and then showing a clip of a typical performance would make it much easier to follow the study designs for scholars. Videos of the conditions bees failed at would be highly useful in order to compare different hypotheses for how the bees are solving this problem. I also think it is highly important to code the videos for switching behaviours. When solving the connected vs unconnected string tasks, when bees were observed pulling the unconnected string, did they quickly switch to the other string? Or did they continue to pull the wrong string? This would help discriminate the use of perceptual-motor feedback from other hypotheses.

We added the test videos as suggested by the reviewer, and we added the data for each bee's choice. However, both connected and disconnected strings were glued to the floor, and therefore perceptual-motor feedback was equal and irrelevant between the choices during the test.

The experiments are also not described well, for my below comments I have assumed that different groups of bees were tested for experiments 1-8, and that experiment 6 was run as described in line 331, where bees were given string-pulling training without perceptual feedback rather than how it is described in Figure 4B, which describes bees as receiving string pulling training with feedback.

We now added figures of Experiment 6 and 7 in the Figure 1B, and we mentioned that different groups of bees were tested for Experiments 1-9.

The authors suggest the bees' performance is best explained by what they term 'image matching'. However, experiment 6 does not seem to support this without assuming retroactive image matching after the problem is solved. The logic of experiment 6 is described as "This was to ensure that the bees could not see the familiar "lollipop shape"

while pulling strings....If the bees prefer to pull the connected strings, this would indicate that bees memorize the arrangement of strings-connected flowers in this task." I disagree with this second sentence, removing perceptual feedback during training would prevent bees memorising the lollipop shape, because, while solving the task, they don't actually see a string connected to a yellow flower, due to the black barrier. At the end of the task, the string is now behind the bee, so unless the bee is turning around and encoding this object retrospectively as the image to match, it seems hard to imagine how the bee learns the lollipop shape.

We agree with the reviewer that while solving the task in the last step during training, the bees don't actually see a string connected to a yellow flower, due to the black barrier. Since the full shape is only visible after the pulling is completed and this requires the bee to "check back" on the entire display after feeding, to basically conclude "this is the shape that I need to be looking for later".

Another possibility is that bumblebees might remember the image of the "lollipop shape" while training the bees in the first step, in which the "lollipop shape" was directly presented to the bumblebee in the early step of the training.

We added the experiment suggested by the reviewer, and the result showed that when a green table was placed behind the string to obscure the "lollipop shape" at any point during the training phase, the bees were unable to identify the connected string. The result further supports that bumblebees learn to choose the connected string through image matching.

Despite this, the authors go on to describe image matching as one of their main findings. For this claim, I would suggest the authors run another experiment, identical to experiment 6 but with a black panel behind the bee, such that the string the bee pulls behind itself disappears from view. There is now no image to match at any point from the bee's perspective so it should now fail the connectivity task.

Strengths:

Despite these issues, this is a fascinating dataset. Experiments 1 and 2 show that the bees are not learning to discriminate between connected and unconnected stimuli rapidly in the first trials of the test. Instead, it is clear that experience in string pulling is needed to discriminate between connected and unconnected strings. What aspect of this experience is important? Experiment 6 suggests it is not image matching (when no image is provided during problem-solving, but only afterward, bees still attend to string connectivity) and casts doubt on perceptual-motor feedback (unless from the bee's perspective, they do actually get feedback that pulling the string moves the flower, video is needed here). Experiments 7 and 8 rule out means-end understanding because if the bees are capable of imagining the effect of their actions on the string and then planning out their actions (as hypotheses such as insight, means-end understanding and string connectivity suggest), they should solve these tasks. If the authors can compare the bees' performance in a more detailed way to other species, and run the experiment suggested, this will be a highly exciting paper

We appreciate the valuable comment from the reviewer. We compared the bees' performance to other species, and conducted the experiment as suggested by the reviewer.

Recommendations for the authors:

Reviewer #1 (Recommendations For The Authors):

Smaller comments:

Line 64: is the word 'simple' needed here? It could also be explained by more complex forms of associative learning, no?

We deleted “simple”.

Methods:

Line 230: was it checked that this was high-contrast for the bees?

We added the relevant reference in the revised manuscript.

Line 240: how much sucrose solution was present in the flowers?

We added 25 microliters sucrose solution in the flowers. We added the information in the revised manuscript.

Line 266: check grammar.

We checked the grammar as follows: “During tests, both strings were glued to the floor of the arena to prevent the air flow generated by flying bumblebees’ wings from changing the position of the string.”

Statistical analysis:

- What does it mean that "Bees identity and colony were analyzed with likelihood ratio tests"?

Bees identity and colony was set as a random variable. We changed the analysis methods in the revised manuscript, and results of the all the experiments did not changed.

- Line 359: do you mean proportion rather than percentage?

We mean the percentage.

- "the number of total choices as weights" - this should be explained further. This is the number of choices that each bee made? What was the variation and mean of this number? If bees varied a lot in this metric, it might make more sense to analyze their first choice (as I see you've done) and their first 10 choices or something like that - for consistency.

This refers to the total number of choices made by each bumblebee. We added the mean and standard error of each bee’s number of choices in Table 1. Some bees pulled the string fewer than 10 times; we chose to include all choices made by each bee.

- More generally I think the first test is more informative than the subsequent choices, since every choice after their first could be affected by feedback they are getting in that test phase. Or rather, they are telling you different things.

All the bees were tested only once, however, you might be referring to the first choice. We used Chi-square test to analyze the bumblebees’ first choices in the test. It is worth noting that both connected and disconnected strings were glued to the floor. The bees were unable to move either the connected or disconnected strings during the tests, and only attempted to pull them. Therefore, the feedback from pulling either the connected or disconnected strings is the same.

- Line 362: *I think I know what you mean, but this should be re-phrased because the "number of" sounds more appropriate for a Poisson distribution. I think what you are testing is whether each individual bee chose the connected or the disconnected string - i.e. a 0 or 1 response for each bee?*

We agree with the reviewer that each bee chose the connected or the disconnected string - i.e. a 0 or 1 response for each bee, but not the number. We clarify this as: "The total number of the choices made by each bee was set as weights."

- Line 364-365: *here and elsewhere, every time you mention a model, make it clear what the dependent and independent variables are. i.e. for the mixed model, the 'bee' is the random factor? Or also the colony that the bee came from? Were these nested etc?*

We clarify this in the revised manuscript. The bee identity and colony is the random factor in the mixed model.

- Line 368: *"Latency to the first choice of each bee was recorded" - why? What were the hypotheses/ predictions here?*

The latency to the first choice was intended to see if the bumblebees were familiarizing with the testing pattern. A shorter delay time might indicate that the bumblebees were more familiar with the pattern.

- Line 371: *"Multiple comparisons among experiments were.." - do you mean 'within' experiments? It seems that treatments should not be compared between different experiments.*

We mean multiple comparisons among different experiments; we clarify this in the revised manuscript.

Results

Experiment 1: From the methods, it sounded like you both analyzed the bees' first choice and their total no. of choices, but in the results section (and Figure 1) I only see the data for all choices combined here.

In table 1 and in the text you report the number of bees that chose each option on their first choice, but there are no statistical results associated with these results. At the very least, a chi square or binomial test could be run.

Line 138: "Interestingly, ten out of fifteen bees pulled the connected string in their first choice" - this is presented like it is a significant majority of bees, but a chi-square test of 10 vs 5 has a p-value = 0.1967

We used the Chi square test to analyzed of the bees' first choice. We also added the analyzed data in the Table 1.

Line 143: "It makes sense because the bees could see the "lollipop shape" once they pulled it out from the table." - this feels more like interpretation (i.e. Discussion) rather than results.

We moved the sentence to the discussion.

Line 162: again this feels more like interpretation/ conjecture than results.

We removed the sentence in the results.

Line 184: check grammar.

We checked the grammar. We changed “task” to “tasks”.

Figures

I really appreciated the overview in Figure 5 - though I think this should be Figure 1? Even if the methods come later in eLife, I think it would be nice to have that cited earlier on (e.g. at the start of the results) to draw the reader's attention to it quickly, since it's so helpful. It also then makes the images at the bottom of what is currently Figure 1 make more sense. I also think that the authors could make it clearer in Figure 5 which strings are connected vs disconnected in the figure (even if it means exaggerating the distance more than it was in real life). I had to zoom in quite a bit to see which were connected vs. not. Alternatively, you could have an arrow to the string with the words "connected" "disconnected" the first time you draw it - and similar labels for the other string conditions.

We appreciate the valuable comment from the reviewer. We changed Figure 5 to Figure 2, and Figure 4 to Figure 1. We cited the Figures at the start of the results. We also changed the gap distance between the disconnected strings. Additionally, we added arrows to indicate “connected” and “disconnected” strings in the Figure.

Figure 1 - I think you could make it clearer that the bars refer to experiments (e.g. have an x-axis with this as a label). Also, check the grammar of the y-axis.

We added the experiments number in the Figures. Additionally, we checked the grammar of the y-axis. We changed “percentages” to “parentage”.

I also think it's really helpful to see the supplementary videos but I think it would be nice to see some examples of the test phase, and not just the training examples.

We added Supplementary videos of the testing phase.

Reviewer #2 (Recommendations For The Authors):

Below are also some minor comments:

L40: "approaches".

We changed “approach” to “approaches”.

L42: but likely mainly due to sampling bias of mammals and birds.

We changed the sentence as follows: String pulling is one of the most extensively used approaches in comparative psychology to evaluate the understanding of causal relationships (Jacobs & Osvath, 2015), with most research focused on mammals and birds, where a food item is visible to the animal but accessible only by pulling on a string attached to the reward (Taylor, 2010; Range et al., 2012; Jacobs & Osvath, 2015; Wakonig et al., 2021).

L64: remove "in this study"

We removed “in this study”.

| L64: *simple associative learning of what? Isn't your image matching associative too?*

We removed “simple”.

| L97: *remove "a" before "connected".*

We removed “a” before “connected”.

| L136-138: *but maybe they could still feel the weight of the flower when pulling?*

Because both strings were glued to the floor in the test phase, the feedback was the same and therefore irrelevant. This information is noted in the General Methods.

| L161: *what are these numbers?*

We removed the latency in the revised manuscript.

| L167/ Table 1: *I realise that the authors never tried slanted strings to check if bumblebees used proximity as a cue. Why?*

This was simply because we wanted to focus on whether bumblebees could recognize the connectivity of the string.

| *Discussion: Why did you only control for colour of the string? What if you had used strings with different textures or smells? Unclear if the authors controlled for "bumblebee smell" on the strings, i.e., after a bee had used the string, was the string replaced by a new one or was the same one used multiple times?*

We used different colors to investigate featural generalization of the visual display of the string connected to the flower in this task. We controlled for color because it is a feature that bumblebees can easily distinguish.

Both the flowers and the strings were used only once, to prevent the use of chemosensory cues. We clarify this in the revised manuscript.

| L182: *since what?*

We deleted “since” in the revised manuscript.

| L182-188: *might be worth mentioning that some crows and parrots known for complex cognition perform poorly on broken strings (e.g., <https://doi.org/10.1098/rspb.2012.1998> ; <https://doi.org/10.1163/1568539X-00003511> ; <https://doi.org/10.1038/s41598-021-94879-x>) and Australian magpies use trial and error (<https://doi.org/10.1007/s00265-023-03326-6>).*

We added the following sentences as suggested by the reviewer: “It is worth noting that some crows and parrots known for complex cognition perform poorly on the broken string task without perceptual feedback or learning. For example, New Caledonian crows use perceptual feedback strategies to solve the broken string-pulling task, and no individual showed a significant preference for the connected string when perceptual feedback was restricted (Taylor et al., 2012). Some Australian magpies and African grey parrots can solve the broken string task, but they required a high number of trials, indicating that learning plays a crucial role in solving this task (Molina et al., 2019; Johnsson et al., 2023).”

| L193: *maybe expand on this to put the task into a natural context?*

We added the following sentences as suggested by the reviewer:

“Different flower species offer varying profitability in terms of nectar and pollen to bumblebees; they need to make careful choices and learn to use floral cues to predict rewards (Chittka, 2017). Bumblebees can easily learn visual patterns and shapes of flower (Meyer-Rochow, 2019); they can detect stimuli and discriminate between differently coloured stimuli when presented as briefly as 25 ms (Nityananda et al., 2014). In contrast, causal reasoning involves understanding and responding to causal relationships. Bumblebees might favor, or be limited to, a visual approach, likely due to the efficiency and simplicity of processing visual cues to solve the string-pulling task. ”

| L204: *is causal understanding the same as means-end understanding?*

Means-end understanding is expressed as goal-directed behavior, which involves the deliberate and planned execution of a sequence of steps to achieve a goal. Includes some understanding of the causal relationship (Jacobs & Osvath, 2015; Ortiz et al., 2019). .

| L235: *this is a very big span of time. Why not control for motivation? Cognitive performance can vary significantly across the day (at least in humans).*

Bumblebee motivation is understood to be rather consistent, as those that were trained and tested came to the flight arena of their own volition and were foragers looking to fill their crop load each time to return it to the colony.

| L232: *what is "(w/w)" ? This occurs throughout the manuscript.*

“w/w” represents the weight-to-weight percentage of sugar.

| L250: *this sentence sounds odd. "containing in the central well.." ?? Perhaps rephrase? Unclear what central well refers to? Did the flowers have multiple wells?*

We rephrased the sentence as follows: For each experiment, bumblebees were trained to retrieve a flower with an inverted Eppendorf cap at the center, containing 25 microliters of 50% sucrose solution, from underneath a transparent acrylic table

| L268: *why euthanise?*

The reason for euthanizing the bees is that new foragers will typically only become active after the current ones were removed from the hive.

| L270: *chemosensory cues answer my concern above. Maybe make it clear earlier.*

We moved this sentence earlier in the result.

| L273: *did different individuals use different pulling strategies? Do you have the data to analyse this? This has been done on birds and would offer a nice comparison.*

We analyzed the string-pulling strategies among different individuals, and provided Supplementary Table 1 to display the performances of each individual in different string-pulling experiments.

| L365: unclear why both models. Would be nice to see a GLM output table.

The duration of pulling different kinds of strings were first tested with the Shapiro-Wilk test to assess data normality. The duration data that conforms to a normal distribution was compared using linear mixed-effects models (LMM), while the data that deviates from normality were examined with a generalized linear-mixed model (GLMM). We added a GLM and GLMM output table in the revised manuscript.

| L377: should be a space between the "." and "This".

We added a space between the "." and "This".

| L383-390: some commas and semicolons are in the wrong places.

We carefully checked the commas and semicolons in this sentence.

Reviewer #3 (Recommendations For The Authors):

Minor comments

Line 32: seems to be missing a word, suggest "the bumblebees' ability to distinguish".

we added "the" in the revised manuscript.

Line 47: it would be good to reference other scholars here, this is the central focus of all work in comparative psychology.

We added the reference in the revised manuscript.

Line 50-61: I think the string-pulling literature could be described in more detail here, with mention of perceptual-motor feedback loops as a competing hypothesis to means-end understanding (see Taylor et al 2010, 2012). It seems a stretch to suggest that "String-pulling studies have directly tested means-end comprehension in various species", when perceptual-motor feedback is a competing hypothesis that we have positive evidence for in several species.

We mentioned the perceptual-motor feedback in the introduction as follow:

"Multiple mechanisms can be involved in the string-pulling task, including the proximity principle, perceptual feedback and means-end understanding (Taylor et al., 2012; Wasserman et al., 2013; Jacobs & Osvath, 2015; Wang et al., 2020). The principle of proximity refers to animals preferring to pull the reward that is closest to them (Jacobs & Osvath, 2015). Taylor et al. (2012) proposed that the success of New Caledonian crows in string-pulling tasks is based on a perceptual-motor feedback loop, where the reward gradually moves closer to the animal as they pull the strings. If the visual signal of the reward approaching is restricted, crows with no prior string-pulling experience are unable to solve the broken string task (Taylor et al., 2012).

However, when a green table was placed behind the string to obscure the "lollipop" structure during the training, the bees could not see the "lollipop" during the initial training stage or after pulling the string from under the table. In this situation, the bees were unable to identify the connected string, further proving that bumblebees chose the connected string based on image matching.

Line 68: suggest remove 'meticulously'.

We removed “meticulously”.

Line 99: This is an exciting finding, can the authors please provide a video of a bee solving this task on its first trial?

We added videos in the supplementary materials.

Line 133: perceptual-motor feedback loops should be introduced in the introduction.

We introduced perceptual-motor feedback loops in the revised manuscript.

Line 136: please clarify the prior experience of these bees, it is not clear from the text.

We clarified the prior experience of these bees as follow: Bumblebees were initially attracted to feed on yellow artificial flowers, and then trained with transparent tables covered by black tape (S7 video) through a four-step process.

Line 138: from the video it is not possible to see the bee's perspective of this occlusion. Do the authors have a video or image showing the feedback the bees received? I think this is highly important if they wish to argue that this condition prevents the use of both image matching and a perceptual-motor feedback loop.

We prevented the use of image matching: the bees were unable to see the flower moving towards them above the table during the training phase in this condition. But the bees may receive visual image both after pulling the string out from the table and in the initial stages of training in this condition.

Line 147: please clarify what experience these bees had before this test.

We added the prior experience of bumblebees before training as follow: We therefore designed further experiments based on Taylor et al. (2012) to test this hypothesis. Bumblebees were first trained to feed on yellow artificial, and then trained with the same procedure as Experiment 2, but the connected strings were coiled in the test.

Line 155: This is a highly similar test to that used in Taylor et al 2012, have the authors seen this study?

We mentioned the reference in the revised manuscript as follows: We therefore designed further experiments based on Taylor et al. (2012) to test this hypothesis.

Line 183: This sentence needs rewriting "Since the vast majority of animals, including dogs 183 (Osthaus et al., 2005), cats (Whitt et al., 2009), western scrub-jays (Hofmann et al., 2016) and azure-winged magpies (Wang et al., 2019) are failing in such tasks spontaneously".

We changed the sentence as suggested by the reviewer as follow: Some animals, including dogs (Osthaus et al., 2005), cats (Whitt et al., 2009), western scrub-jays (Hofmann et al., 2016) and azure-winged magpies (Wang et al., 2019) fail in such task spontaneously.

Line 186: "complete comprehension of the functionality of strings is rare" I am not sure the evidence in the current literature supports any animal showing full understanding,

| *can the authors explain how they reach this conclusion?*

We wished to say that few animal species could distinguish between connected and disconnected strings without trial and error learning. We revised the sentence as follows:

It is worth noting that some crows and parrots known for complex cognition perform poorly on broken string task without perceptual feedback or learning. For example, New Caledonian crows use perceptual feedback strategies to solve broken string-pulling task, and no individual showed a significant preference for the connected string when perceptual feedback is restricted (Taylor et al., 2012). Some Australian magpies and African grey parrots can solve the broken string task, but it required a high number of trials, indicating that learning plays a crucial role in solving this task (Molina et al., 2019; Johnsson et al., 2023).

| *Line 190: the authors need to clarify which part of their study provides positive evidence for this conclusion.*

We added the evidence for this conclusion as follows: Our findings suggest that bumblebees with experience of string pulling prefer the connected strings, but they failed to identify the interrupted strings when the string was coiled in the test.

| *Line 265: was the far end of the string glued only?*

The entire string was glued to the floor, not just the far ends of the string.

<https://doi.org/10.7554/eLife.97018.2.sa0>