Supplemental Materials

**Table S1**

*Examples of process and general praise*

|  |  |
| --- | --- |
| Praise Category | |
| Process Praise  (Experiments 1 & 2) | General Praise  (Experiment 2 Only) |
| “Good job!”  “You worked so hard!”  “Good spinning!”  “You’re trying!”  “You’re getting it!”  “Good going!” | “Yay!”  “Great!”  “Right on!”  “That’s neat!”  “Cool!”  “That’s right!” |

Chart, bar chart

Description automatically generated

**Figure S1.** Distribution of praise overlap categories for process praise in Experiment 1 and Experiment 2. Bars represent proportions of total process praise.

**Table S2**

*Correlations between all predictor variables, controlling for multiple comparisons*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Experiment 1 | | | | |
|  | Trying | Success | Both | Random | Overall |
| Trying | – | .28 | .26 | -.11 | .39 |
| Success |  | – | .48† | -.14 | .39 |
| Both |  |  | – | .24 | .79\*\*\* |
| Random |  |  |  | – | .70\*\*\* |
| Overall |  |  |  |  | – |
|  | Experiment 2 | | | | |
|  | Trying | Success | Both | Random | Overall |
| Trying | – | .12 | .41\*\* | .69\*\*\* | .74\*\*\* |
| Success |  | – | .32† | .12 | .26 |
| Both |  |  | – | .37\* | .78\*\*\* |
| Random |  |  |  | – | .83\*\*\* |
| Overall |  |  |  |  | – |

Note: †*p* < .07, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001

**Alternative Measures of Persistence**

In the main text, we utilize Gear Stacking Rate as our dependent measure of infants’ persistence, including both independent gear placements and those assisted by caregivers. In this way, we generate a measure less directly linked to timing than overall trying time and we can capture times when children are persisting but may lack the motoric skills to generate success. However, this measure may call into question whether more hands-on caregivers (i.e., those who assist more with placements) are also caregivers that provide more temporally aligned process praise by nature of their engagement. Thus, we also created a composite measure of infant persistence by summing the z-scores of Gear Stacking Rate and Trying Time for each participant. The pattern of results is the same using this measure. Likewise, when timing effects are analysed using only infants’ independent stacking rates (excluding caregiver-assisted placements), the pattern of results is extremely similar.

**Table S3**

*Models predicting alternative measure of persistence from all process praise overlap categories*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models Predicting Composite Persistence Score | | | | |
| Experiment 1 | |  | Experiment 2 | |
| Predictors | β (SE) | [95% CI] |  | β (SE) | [95% CI] |
| Intercept  Age  Sqrt(Trying)  Sqrt(Success)  Sqrt(Both)  Sqrt(Random) | -1.23 (29.95)  0.05 (1.66)  0.01 (0.45)  0.72 (0.77)  1.07 (0.44)\*  -0.44 (0.32) | [-59.92, 57.46]  [-3.20, 3.29]  [-0.88, 0.90]  [-0.80, 2.24]  [ 0.20, 1.95]  [-1.07, 0.19] |  | -2.15 (1.14)  0.10 (0.05)\*  -0.20 (0.53)  -0.09 (0.91)  1.13 (0.30)\*\*\*  -0.97 (0.36)\* | [-4.37, 0.08]  [ 0.004, 0.19]  [-1.24, 0.84]  [-1.87, 1.68]  [ 0.54, 1.73]  [-1.68, -0.26] |
|  | Models Predicting Independent Gear Stacking Rate | | | | |
| Experiment 1 | |  | Experiment 2 | |
| Predictors | β (SE) | [95% CI] |  | β (SE) | [95% CI] |
| Intercept  Age  Sqrt(Trying)  Sqrt(Success)  Sqrt(Both)  Sqrt(Random) | -5.96 (10.86)  0.36 (0.60)  -0.13 (0.16)  -0.17 (0.28)  0.43 (0.16)\*  -0.14 (0.12) | [-27.24, 15.33]  [-0.82, 1.54]  [-0.46, 0.19]  [-0.72, 0.38]  [ 0.11, 0.74]  [-0.36, 0.09] |  | -0.61 (0.80)  0.10 (0.03)\*\*  -0.53 (0.38)  -0.32 (0.64)  0.66 (0.21)\*\*  -0.32 (0.26) | [-2.18, 0.96]  [ 0.03, 0.16]  [-1.26, 0.21]  [-1.58, 0.93]  [ 0.24, 1.07]  [-0.82, 0.18] |

Note: \**p* < .05, \*\**p* < .01, \*\*\**p* < .001

**Table S4**

*Models predicting alternative measure of persistence from Both and overall process praise*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Models Predicting Composite Persistence Score | | | | |
| Experiment 1 | |  | Experiment 2 | |
| Predictors | β (SE) | [95% CI] |  | β (SE) | [95% CI] |
| Intercept  Age  Sqrt(Overall)  Sqrt(Both) | 6.58 (30.35)  -0.39 (1.68)  -0.19 (0.35)  1.23 (0.49)\* | [-52.90, 66.05]  [-3.68, 2.90]  [-0.88, 0.50]  [0.29, 2.22] |  | -1.92 (1.06)  0.09 (0.04)\*  -0.96 (0.28)\*\*  1.81 (0.38)\*\*\* | [-4.00, 0.15]  [0.005, 0.18]  [-1.51, -0.42]  [1.06, 2.56] |
|  | Models Predicting Independent Gear Stacking Rate | | | | |
| Experiment 1 | |  | Experiment 2 | |
| Predictors | β (SE) | [95% CI] |  | β (SE) | [95% CI] |
| Intercept  Age  Sqrt(Overall)  Sqrt(Both) | -6.74 (10.33)  0.41 (0.57)  -0.17 (0.12)  0.45 (0.17)\* | [-26.99, 13.51]  [-0.71, 1.52]  [-0.40, 0.07]  [0.12, 0.78] |  | -0.73 (0.75)  0.10 (0.03)\*\*  -0.57 (0.20)\*\*  0.96 (0.27)\*\*\* | [-2.20, 0.75]  [0.04, 0.17]  [-0.95, -0.18]  [0.42, 1.50] |

Note: \**p* < .05, \*\**p* < .01, \*\*\**p* < .001

**Table S5**

*Models predicting gear stacking rate from Both and overall process praise*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Experiment 1 | |  | Experiment 2 | |
| Predictors | β (SE) | [95% CI] |  | β (SE) | [95% CI] |
| Intercept  Age  Sqrt(Overall)  Sqrt(Both) | -5.27 (10.41)  0.33 (0.58)  -0.08 (0.12)  0.46 (0.17)\* | [-25.67, 15.15]  [-0.80, 1.46]  [-0.31, 0.16]  [0.13, 0.79] |  | -0.88 (0.75)  0.13 (0.03)\*\*\*  -0.69 (0.20)\*\*\*  1.23 (0.27)\*\*\* | [-2.35, 0.58]  [0.07, 0.19]  [-1.07, -0.30]  [0.70, 1.77] |

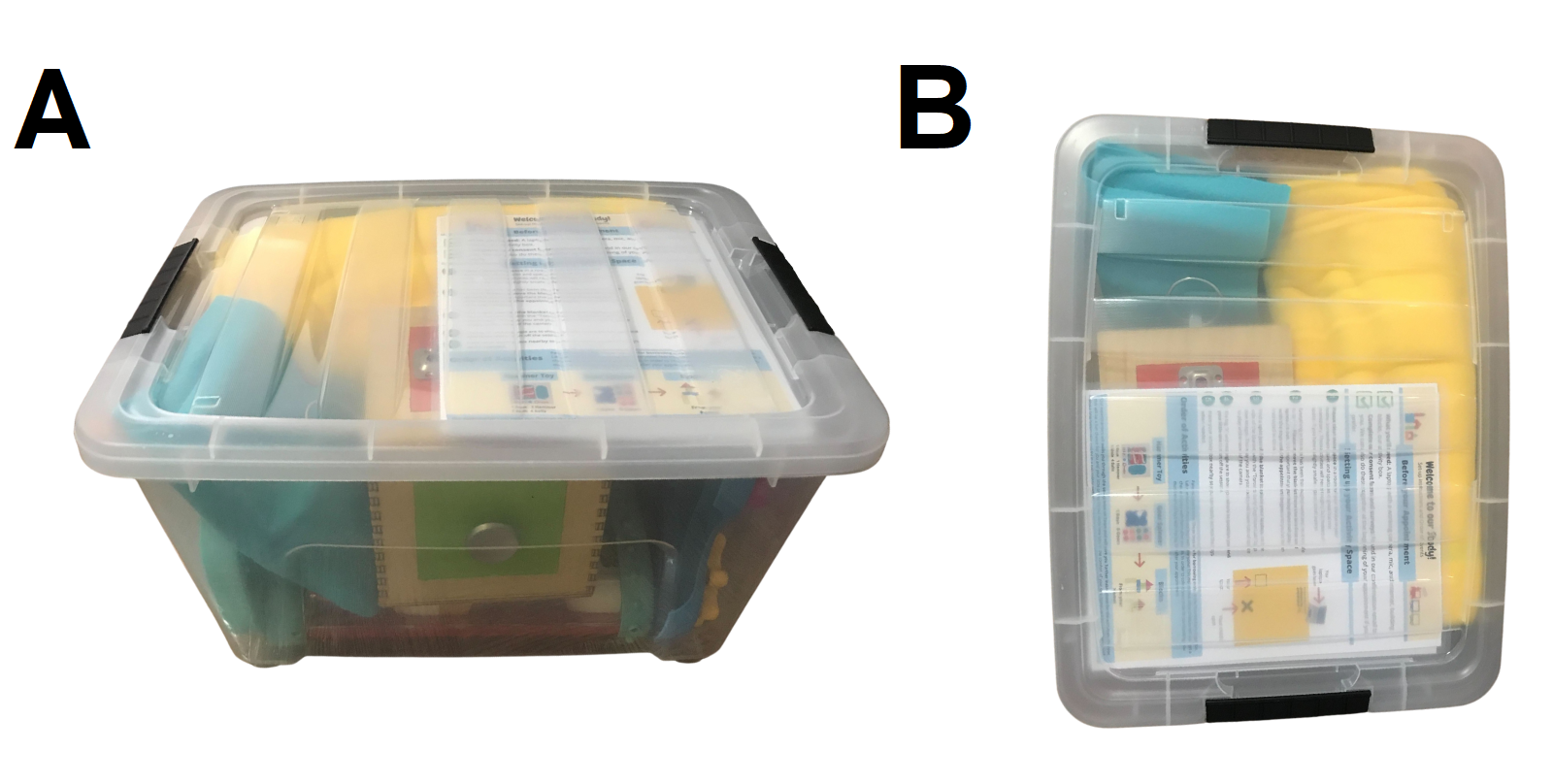
Note: \**p* < .05, \*\**p* < .01, \*\*\**p* < .001

**Additional Information for Experiment 2 Methods**

*Delivery of Materials.* As this experiment was conducted during the COVID-19 pandemic, a contactless delivery procedure was used to provide caregivers with all necessary materials for the study session. To facilitate experimental control, materials were delivered in identical study kits and always arranged the same way. Families were provided with a yellow blanket with markings designating where the infant and caregiver should sit, the gear stacking toy, and materials for two additional experiments run during the same session.

Before participating in the experimental session, caregivers first met with an experimenter over Zoom who explained briefly how caregivers should prepare for the session and to coordinate details for delivery of the study kit. Caregivers were encouraged to prepare a room with minimized distractions (e.g., pets, other family members) with enough space to spread out the blanket in the study kit, and to not open the kit or look at the materials until the experimental session could be conducted.

Once kits were delivered to families’ doors, experimenters guided the caregivers to take out the relevant study materials over Zoom and asked caregivers to conceal the kit and the remainder of its contents from infants until further instruction. After the session was completed, kits were collected from families and sanitized before delivery for the next experimental session.



**Figure S2.** Images of study kits delivered to families’ homes in Experiment 2. Materials were always arranged identically. (A) Shows a lateral view, and (B) shows a top-down view.

*Warm-Up.* The warm-up task was intended to also serve as a measure of independent persistence for both Experiment 2 as well as a separate experiment investigating the effects of caregiver prompting on infant persistence. As part of this task, caregivers were to deliver prompts when experimenters flashed instructions on the screen using Zoom screenshare. However, analyses of this task are excluded here as caregiver compliance was generally low. Specifically, 53% of caregivers delivered an incorrect number of prompts, and among the prompts that were delivered only 77% of prompts were delivered correctly. Thus, coding revealed issues with both the quantity and quality of the prompts delivered, this task was utilized only to establish an exclusion criterion. Caregivers who delivered 8 or more prompts incorrectly were excluded from the final sample. Thus, the warm-up task results in several exclusions (*n* = 8), both for failure to engage for at least 2 minutes (*n* = 5) and caregiver compliance (*n* = 3).

**Descriptive Analyses of General Praise in Experiment 2**

We sought to descriptively assess caregivers’ general praise overlap with trying and success in a similar manner to the analyses of process praise overlap in the main text. Caregivers offered a total of 403 utterances containing general praise (see Table S6 & Figure S3). Of these utterances: 61 were classified as Trying (15%), 40 as Success (10%), 172 as Both (43%), and 130 as Random (32%). To examine whether general praise was more likely to be temporally aligned to infant behavior than not, we conducted a binomial test comparing the total number of general praise utterances which overlapped with key times to a chance reference (i.e., 50%). Just as we found with process praise in Experiment 2, we observed evidence that most general praise occurred when toddlers were trying and/or succeeding in the (*p* < .001).

**Table S6**

*Counts of coded temporal overlap categories by praise type*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Overlap Period | | | |  |
| Praise Type | Trying | Success | Both | Random | Total |
| Process Praise | 30 | 7 | 75 | 72 | 184 |
| General Praise | 61 | 40 | 172 | 130 | 403 |
| Total | 91 | 47 | 247 | 202 | 587 |

Chart, bar chart

Description automatically generated

**Figure S3.** Comparison of overlap categories coded in Experiment 2, divided by general and process praise. Bars represent the category as a proportion of the overall praise within type (e.g., General Both represents the proportion of general praise that occurred during both Trying & Success).

We further sought to understand whether there were descriptive differences between process praise and general praise overlap. To this end, a chi-square test of independence was performed to examine the relation praise type and overlap category. The relation between these variables was significant,  *x*2 (3, *N* = 587) = 7.87, *p* = .05. This suggested that the praise categories were not uniformly distributed across praise type and overlap time. Given the aims of our experiments, we followed up this finding with five exploratory contrasts. First, to understand whether there was a difference in the overall quantity of process praise and general praise, and second, we sought to understand whether there were differences in temporal alignment as a function of praise type. As such, we utilized a set of exploratory contrasts comparing the proportion of each overlap category between praise types (e.g., comparing the proportion of Trying process praise to Trying general praise).

Thus, we utilized a binomial test to compare the overall proportion of praise which was general praise to chance (i.e., 0.5). This analysis revealed that caregivers utilized significantly more general praise overall (69%) than process praise (31%, *p* < .001). This result likely reflected the difference in the specificity of the two types of praise – while process praise requires specific verbal reference to children’s actions, general praise is given more automatically and is sometimes used as a positive filler word. When we further performed two-sample tests for equality of proportions we did not find evidence that the proportion of praise that was coded as Trying, Both, or Random differed between process praise and general praise (all *p’s* ≥ .13). However, it did appear that a greater proportion of general praise was coded as Success (10%) relative to process praise (4%; *p* = .01). While we did not have *a priori* hypotheses about differences between these praise types, it may be the case that because of the casual nature of general praise, that caregivers offered it more reflexively when gears were spinning down the rod.

**Table S7**

*Correlations between process praise and general praise overlap categories, coded as proportions of overall praise*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Overlap Times | | | | | | | |
|  |  | Process | | | | General | | | |
| Type | Time | Trying | Success | Both | Random | Trying | Success | Both | Random |
| Process | Trying | – | 0.26† | 0.19 | 0.40\*\* | -0.16 | 0.27\* | -0.18 | 0.11 |
| Success |  | – | 0.31\* | 0.16 | -0.15 | 0.34\* | -0.05 | -0.04 |
| Both |  |  | – | 0.04 | -0.20 | 0.39\*\* | 0.09 | -0.22 |
| Random |  |  |  | – | -0.04 | 0.09 | -0.39\*\* | -0.08 |
| General | Trying |  |  |  |  | – | -0.26† | -0.16 | -0.21 |
| Success |  |  |  |  |  | – | 0.09 | -0.14 |
| Both |  |  |  |  |  |  | – | -0.30\* |
| Random |  |  |  |  |  |  |  | – |