**Do Children Use Language Structure to Discover the Recursive Rules of Counting?**

Schneider, Sullivan, Marušič, Žaucer, Biswas, Mišmaš, Plesničar, & Barner

**Supplementary Online Material**

**Table of Contents**

1. **Unit Task: Full model details and interim model results**
2. **Next Number Task: Full model details and interim model results**
3. **Analyses of <100 and >100 items**
4. **Items beyond Highest Count**
5. **Mean performance by Counting Resilience**
6. **Highest Contiguous Next Number**
7. **Initial Highest Count**
8. **Types/Frequency of counting errors**
9. **Unit Task: Model building process and interim results**

We constructed Unit Task generalized linear mixed effects models using the ‘lme4’ package in R (Bates et al., 2015). The base model had the following formula: Correct ~ Within/Outside IHC + Item Magnitude + Age + (1|Subject). Continuous variables were scaled and centered to facilitate model fit. We first constructed four individual models with the following candidate measures of productivity: (1) Counting Resilience; (2) Final Highest Count; (3) Initial Highest Count; and (4) Highest Contiguous Next Number. Within each language, we conducted a Likelihood Ratio Test between each of Models 1-4 and the base model to determine whether a candidate productivity measure was significantly related to Unit Task performance.

As preregistered, after determining which productivity measures were individually predictive of successor knowledge, we constructed a single large model within each language using hierarchical model comparison. For each language we built a large model containing the productivity measure associated with the lowest AIC. We then added the other productivity measures which significantly predicted Unit Task performance in that language in order of increasing AIC. We performed a Likelihood Ratio Test with the addition of each new term, and retained that term on the basis of a significant χ2 statistic.

**1.1. Cantonese.**Our individual models indicated three productivity measures significantly improved the fit of the base model (Table 1): Final Highest Count (χ2(1) = 11.66, *p* = .0006); Initial Highest Count (χ2(1) = 38.91, *p* < .0001); and Highest Contiguous Next Number (χ2(1) = 11.83, *p* = .0006). Resilience was not significantly related to Unit Task performance (χ2(1) = 0.05, *p* = .82). The base for our large model contained Initial Highest Count, which yielded the lowest AIC in our individual Likelihood Ratio Tests. Adding Highest Contiguous Next Number did not significantly improve the fit of this model (χ2(1) = 1.86, *p* = .17), nor did Final Highest Count (χ2(1) = 1.39, *p* = .24).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Cantonese** |  | | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | | Model 2: FHC | Model 3: IHC | Model 4: HCNN | |
| (Intercept) | 0.53\*\*\* | 0.51\*\* | | 0.58\*\*\* | 0.69\*\*\* | 0.56\*\*\* | |
| Resilient | — | 0.05 | | — | — | — | |
| FHC | — | — | | 0.44\*\* | — | — | |
| IHC | — | — | | — | 0.81\*\*\* | — | |
| HCNN | — | — | | — | — | 0.39\*\* | |
| Trial Within IHC | 0.26 | 0.26 | | 0.15 | -0.05 | 0.21 | |
| Item Magnitude | -0.39\*\*\* | -0.39\*\*\* | | -0.44\*\*\* | -0.52\*\*\* | -0.41\*\*\* | |
| Age | 0.56\*\*\* | 0.55\*\*\* | | 0.26\* | 0 | 0.37\*\*\* | |
| AIC | 1626.2 | 1628.1 | | 1616.5 | 1589.3 | 1616.3 | |
| Conditional R2 | 0.261 | 0.261 | | 0.262 | 0.266 | 0.265 | |

Table 1. Base and individual productivity model regression models for predicting Unit Task performance in Cantonese. Coefficients significance was calculated using the standard normal approximation to the *t* distribution (Barr, Levy, Scheepers, & Tily, 2013); *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001. Conditional R2 calculated using both fixed and random effects (Nakagawa, Johnson, & Schielzeth, 2017).

**1.2. Slovenian.**Our individual models indicated all four productivity measures significantly improved the fit of the base model (Table 2): Resilience (χ2(1) = 15.97, *p* < .0001); Final Highest Count (χ2(1) = 25.13, *p* < .0001); Initial Highest Count (χ2(1) = 14.59, *p* = .0001); and Highest Contiguous Next Number (χ2(1) = 26.23, *p* < .0001). The base for our large model contained Highest Contiguous Next Number, which yielded the lowest AIC in our individual Likelihood Ratio Tests. The addition of Final Highest Count significantly improved the fit of this model (χ2(1) = 8.03, *p* = .005). Adding Resilience to a model containing both Final Highest Count and Highest Contiguous Next Number did not improve its fit (χ2(1) = 0.07, *p* = .80). Finally, Initial Highest Count did not explain unique variance when added to this model (χ2(1) = 0.02, *p* = .89).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Slovenian** |  | | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | | Model 2: FHC | Model 3: IHC | Model 4: HCNN | |
| (Intercept) | 0.15 | -0.08 | | 0.21\* | 0.21\* | 0.19 | |
| Resilient | — | 0.95\*\*\* | | — | — | — | |
| FHC | — | — | | 0.57\*\*\* | — | — | |
| IHC | — | — | | — | 0.44\*\*\* | — | |
| HCNN | — | — | | — | — | 0.53\*\*\* | |
| Trial Within IHC | 0.43\* | 0.35 | | 0.27 | 0.27 | 0.35 | |
| Item Magnitude | -0.28\*\*\* | -0.31\*\*\* | | -0.33\*\*\* | -0.33\*\*\* | -0.31\*\*\* | |
| Age | 0.41\*\*\* | 0.23\* | | 0.12 | 0.25\* | 0.24\*\* | |
| AIC | 1436.7 | 1422.7 | | 1413.6 | 1424.1 | 1412.5 | |
| Conditional R2 | 0.208 | 0.212 | | 0.216 | 0.217 | 0.218 | |

Table 2. Base and individual productivity model regression models for predicting Unit Task performance in Slovenian. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**1.3. English (US).**Our individual models indicated three productivity measures significantly improved the fit of the base model (Table 3): Final Highest Count (χ2(1) = 18.17, *p* < .0001); Initial Highest Count (χ2(1) = 48.78, *p* < .0001); and Highest Contiguous Next Number (χ2(1) = 35.66, *p* < .0001). The base for our large model contained Initial Highest Count, which yielded the lowest AIC in our individual Likelihood Ratio Tests. Adding Highest Contiguous Next Number significantly improved the fit of this model (χ2(1) = 8.65, *p* = .003). The addition of Final Highest Count to a model containing both Highest Contiguous Next Number and Initial Highest Count did not improve the fit of this model (χ2(1) = 0.10, *p* = .75).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **English (US)** |  | | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | | Model 2: FHC | Model 3: IHC | Model 4: HCNN | |
| (Intercept) | 0.43 | 0.32 | | 0.47\*\*\* | 0.57\*\*\* | 0.50\*\*\* | |
| Resilient | — | 0.28 | | — | — | — | |
| FHC | — | — | | 0.64\*\*\* | — | — | |
| IHC | — | — | | — | 1.01\*\*\* | — | |
| HCNN | — | — | | — | — | 0.81\*\*\* | |
| Trial Within IHC | 0.35 | 0.34 | | 0.24 | 0.08 | 0.27 | |
| Item Magnitude | -0.41\*\*\* | -0.41\*\*\* | | -0.44\*\*\* | -0.49\*\*\* | -0.43\*\*\* | |
| Age | 0.69\*\*\* | 0.62\*\*\* | | 0.25 | 0.10 | 0.26\* | |
| AIC | 1510.3 | 1511.2 | | 1494.1 | 1463.5 | 1476.7 | |
| Conditional R2 | 0.333 | 0.334 | | 0.339 | 0.360 | 0.352 | |

Table 3. Base and individual productivity model regression models for predicting Unit Task performance in US English. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**1.4. Hindi.**The results of our individual model comparisons indicated that three candidate measures of productivity significantly improved the base model (Table 4): Final Highest Count (χ2(1) = 21.60, *p* < .0001); Initial Highest Count (χ2(1) = 28.36, *p* < .0001); and Highest Contiguous Next Number (χ2(1) = 27.63, *p* < .0001). The base for our large model contained Initial Highest Count, which was associated with the lowest AIC in our individual model comparisons. The addition of Highest Contiguous Next Number to this model significantly improved its fit (χ2(1) = 9.91, *p* = .002). The addition of Final Highest Count to a model containing both Initial Highest Count and Highest Contiguous Next Number did not significantly improve its fit (χ2(1) = 0.93, *p* = .34).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Hindi** |  | | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | | Model 2: FHC | Model 3: IHC | Model 4: HCNN | |
| (Intercept) | -0.54\*\*\* | -0.59\*\*\* | | -0.51\*\*\* | -0.49\*\*\* | -0.50\*\*\* | |
| Resilient | — | 0.61 | | — | — | — | |
| FHC | — | — | | 0.45\*\*\* | — | — | |
| IHC | — | — | | — | 0.53\*\*\* | — | |
| HCNN | — | — | | — | — | 0.51\*\*\* | |
| Trial Within IHC | 0.14 | 0.14 | | 0.07 | 0.01 | 0.12 | |
| Item Magnitude | -0.36\*\*\* | -0.36\*\*\* | | -0.38\*\*\* | -0.40\*\*\* | -0.37\*\*\* | |
| Age | 0.50\*\*\* | 0.48\*\*\* | | 0.41\*\*\* | 0.32\*\*\* | 0.37\*\*\* | |
| AIC | 1344.7 | 1343.5 | | 1325.1 | 1318.4 | 1319.1 | |
| Conditional R2 | 0.211 | 0.211 | | 0.213 | 0.215 | 0.214 | |

Table 4. Base and individual productivity model regression models for predicting Unit Task performance in Hindi. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**1.5. Gujarati.**Our individual model comparisons indicated that none of our candidate productivity measures significantly improved the base model (Table 5).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Gujarati** |  | | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | | Model 2: FHC | Model 3: IHC | Model 4: HCNN | |
| (Intercept) | -0.35\*\*\* | -0.38\*\*\* | | -0.34\*\*\* | -0.34\*\*\* | -0.34\*\*\* | |
| Resilient | — | 0.20 | | — | — | — | |
| FHC | — | — | | 0.09 | — | — | |
| IHC | — | — | | — | 0.11 | — | |
| HCNN | — | — | | — | — | 0.11 | |
| Trial Within IHC | 0.54\*\* | 0.53\*\* | | 0.50\* | 0.48\* | 0.51\*\* | |
| Item Magnitude | -0.16 | -0.17 | | -0.17 | -0.18\* | -0.17 | |
| Age | 0.19\* | 0.18\* | | 0.18\* | 0.17\* | 0.18\* | |
| AIC | 1272.7 | 1274.0 | | 1273.7 | 1273.0 | 1273.1 | |
| Conditional R2 | 0.093 | 0.094 | | 0.093 | 0.093 | 0.093 | |

Table 5. Base and individual productivity model regression models for predicting Unit Task performance in Gujarati. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**1.6. Cross-linguistic models.** The results of our individual model comparisons revealed that only Highest Contiguous Next Number significantly improved the base model in both Experiment 1 (Table 6, χ2(1) = 22.70, *p* < .0001) and Experiment 2 (Table 7, χ2(1) = 13.18, *p* =.0003).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Comparison to Cantonese** | | | | **Comparison to Slovenian** | | | | |
|  | ***Coefficient Estimates (β)*** | | | | ***Coefficient Estimates (β)*** | | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC | Model 3: HCNN | | Base | Model 1: Resilience | Model 2: FHC | Model 3: HCNN |
| (Intercept) | 0.21\* | 0.14 | 0.19 | 0.31\*\* | | 0.57\*\*\* | 0.51\*\*\* | 0.57\*\*\* | 0.50\*\*\* |
| Resilient | — | 0.15 | — | — | | — | 0.15 | — | — |
| FHC | — | — | 0.15 | — | | — | — | 0.15 | — |
| HCNN | — | — | — | 0.34\*\*\* | | — | — | — | 0.34\*\*\* |
| Cantonese | — | — | — | — | | -0.35\* | -0.36\* | -0.38\* | -0.19 |
| Slovenian | 0.35\* | 0.36\* | 0.38\* | 0.20 | | — | — | — | — |
| English (US) | 0.60\*\*\* | 0.61\*\*\* | 0.61\*\*\* | 0.43\*\* | | 0.25 | 0.24 | 0.24 | 0.24 |
| IHC | 0.68\*\*\* | 0.67\*\*\* | 0.58\*\*\* | 0.54\*\*\* | | 0.77\*\*\* | 0.72\*\*\* | 0.62\*\* | 0.55\*\* |
| Trial Within IHC | 0.10 | 0.10 | 0.10 | 0.10 | | 0.10 | 0.10 | 010 | 0.10 |
| Item Magnitude | -0.44\*\*\* | -0.44\*\*\* | -0.44\*\*\* | -0.44\*\*\* | | -0.44\*\*\* | -0.44\*\*\* | -0.44\*\*\* | -0.44\*\*\* |
| Age | 0.16 | 0.13 | 0.12 | 0.10 | | 0.16 | 0.13 | 0.12 | 0.10 |
| WPPSI | 0.09 | 0.08 | 0.08 | 0.07 | | 0.09 | 0.08 | 0.08 | 0.07 |
| Cantonese: IHC | — | — | — | — | | -0.10 | -0.06 | -0.04 | -0.01 |
| Slovenian: IHC | 0.10 | 0.06 | 0.04 | 0.01 | | — | — | — | — |
| English (US): IHC | 0.31\* | 0.30\* | 0.30\* | 0.20 | | 0.21 | 0.24 | 0.25 | 0.19 |
| AIC | 4400.02 | 4400.5 | 4400.00 | 4379.3 | |  |  |  |  |
| Conditional R2 | 0.292 | 0.292 | 0.291 | 0.293 | |  |  |  |  |

Table 6. Base and individual productivity model regression models for predicting Unit Task performance in cross-linguistic analyses with Cantonese, Slovenian, and US English, with Cantonese (left) and Slovenian (right) selected as reference groups. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | ***Coefficient Estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC |  | Model 3: HCNN |
| (Intercept) | 0.45\*\*\* | 0.44\*\*\* | 0.45\*\*\* |  | 0.43\*\*\* |
| Resilient | - | 0.02 | - |  | - |
| FHC | - | - | -0.01 |  | - |
| HCNN | - | - | - |  | 0.26\*\*\* |
| Hindi | -0.71\*\*\* | -0.71\*\*\* | -0.71\*\*\* |  | -0.70\*\*\* |
| Gujarati | -0.75\*\*\* | -0.75\*\*\* | -0.75\*\*\* |  | -0.72\*\*\* |
| IHC | 0.59\*\*\* | 0.59\*\*\* | 0.60\*\*\* |  | 0.45\*\*\* |
| Trial Within IHC | 0.15 | 0.15 | 0.15 |  | 0.16 |
| Item Magnitude | -0.37\*\*\* | -0.37\*\*\* | -0.37\*\*\* |  | -0.37\*\*\* |
| Age | 0.30\*\* | 0.30\*\* | 0.30\*\* |  | 0.27\*\* |
| WPPSI | 0.12\* | 0.12\* | 0.12\* |  | 0.10 |
| Hindi: IHC | 0.56\*\* | 0.56\*\* | 0.56\*\* |  | 0.45\* |
| Gujarati: IHC | -0.37\* | -0.38\* | -0.37\* |  | -0.42\* |
| AIC | 3951.63 | 3953.6 | 3953.6 |  | 3940.5 |
| Conditional R2 | 0.281 | 0.281 | 0.281 |  | 0.284 |

Table 7. Base and individual productivity model regression models for predicting Unit Task performance in cross-linguistic analyses with Hindi, Gujarati, and US English, with US English as the reference group. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

1. **Next Number Task: Model building process and interim results**

We constructed our Next Number Task models with the same process we used in our Unit Task analyses. The base model had the following formula: Correct ~ Within/Outside IHC + Item Magnitude + Age + (1|Subject). Continuous variables were scaled and centered to facilitate model fit. We first built three individual models with the following candidate measures of productivity: (1) Counting Resilience; (2) Final Highest Count; and (3) Initial Highest Count.

**2.1. Cantonese.**Individual models indicated two significant predictors of Next Number performance (Table 8): Final Highest Count (χ2(1) = 34.33, *p* < .0001) and Initial Highest Count (χ2(1) = 68.63, *p* < .0001). The base for our large model predicting Next Number performance contained Initial Highest Count, which was associated with the lowest AIC in our individual model comparisons. The addition of Final Highest Count to this model did not significantly improve its fit (χ2(1) = 0.37, *p* = .54).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cantonese** | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC | Model 3: IHC | |
| (Intercept) | -0.69\*\*\* | -0.84\*\* | -0.62\*\*\* | -0.47\*\* | |
| Resilient | — | 0.31 | — | — | |
| FHC | — | — | 1.20\*\*\* | — | |
| IHC | — | — | — | 1.62\*\*\* | |
| Trial Within IHC | 1.20\*\*\* | 1.19\*\*\* | 1.03\*\*\* | 0.75\*\* | |
| Item Magnitude | -0.92\*\*\* | -0.92\*\*\* | -0.99\*\*\* | -1.12\*\*\* | |
| Age | 1.49\*\*\* | 1.44\*\*\* | 0.67\*\*\* | 0.34\* | |
| AIC | 1203.5 | 1204.7 | 1171.2 | 1136.9 | |
| Conditional R2 | 0.674 | 0.674 | 0.678 | 0.677 | |

Table 8. Base and individual productivity model regression models for predicting Next Number Task performance in Cantonese. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**2.2. Slovenian.** Individual models revealed that all three candidate measures were significantly related to Next Number performance (Table 9): Resilience (χ2(1) = 40.26, *p* < .0001); Final Highest Count (χ2(1) = 72.93, *p* < .0001); and Initial Highest Count (χ2(1) = 36.19, *p* < .0001). We constructed the base for our large model with Final Highest Count, which was associated with the lowest AIC in our individual model comparisons. The addition of neither Resilience (χ2(1) = 0.004, *p* = .99) nor Initial Highest Count (χ2(1) = 0.10, *p* = .76) significantly improved the fit of this model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Slovenian** | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC | Model 3: IHC | |
| (Intercept) | -0.63\*\* | -1.35\*\*\* | -0.31\* | -0.44\* | |
| Resilient | — | 2.86\*\*\* | — | — | |
| FHC | — | — | 1.83\*\*\* | — | |
| IHC | — | — | — | 1.37\*\*\* | |
| Trial Within IHC | 1.46\*\*\* | 1.37\*\*\* | 1.19\*\*\* | 1.26\*\*\* | |
| Item Magnitude | -0.95\*\*\* | -1.01\*\*\* | -1.09\*\*\* | -1.04\*\*\* | |
| Age | 1.54\*\*\* | 0.89\*\*\* | 0.47\*\* | 0.97\*\*\* | |
| AIC | 901.56 | 863.31 | 830.63 | 867.38 | |
| Conditional R2 | 0.698 | 0.693 | 0.698 | 0.706 | |

Table 9. Base and individual productivity model regression models for predicting Next Number Task performance in Slovenian. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**2.3. English (US).**Our individual models indicated that all three candidate measures significantly predicted Next Number performance (Table 10): Resilience (χ2(1) = 14.76, *p* = .0001); Final Highest Count (χ2(1) = 44.00, *p* < .0001); and Initial Highest Count (χ2(1) = 30.52, *p* < .0001). We constructed our large model with Final Highest Count, which was associated with the lowest AIC in our individual model comparisons. in the base. The addition of Initial Highest Count to this base only marginally improved the fit of the model (χ2(1) = 3.73, *p* = .053), and Resilience did not explain any additional variance (χ2(1) = 0.70, *p* = .40).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **English (US)** | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC | Model 3: IHC | |
| (Intercept) | -0.50\* | -1.13\*\*\* | -0.42\* | -0.32 | |
| Resilient | — | 1.59\*\*\* | — | — | |
| FHC | — | — | 1.49\*\*\* | — | |
| IHC | — | — | — | 1.33\*\*\* | |
| Trial Within IHC | 1.81\*\*\* | 1.81\*\*\* | 1.64\*\*\* | 1.57\*\*\* | |
| Item Magnitude | -0.59\*\*\* | -0.59\*\*\* | -0.66\*\*\* | -0.69\*\*\* | |
| Age | 1.76\*\*\* | 1.33\*\*\* | 0.69\*\* | 0.99\*\*\* | |
| AIC | 1054.78 | 1041.8 | 1012.6 | 1026.1 | |
| Conditional R2 | 0.716 | 0.713 | 0.718 | 0.737 | |

Table 10. Base and individual productivity model regression models for predicting Next Number Task performance in US English. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**2.4. Hindi.**Individual model comparisons revealed that all three candidate productivity measures significantly improved the fit of the base model (Table 11): Resilience (χ2(1) = 4.55, *p* = .03); Final Highest Count (χ2(1) = 27.38, *p* < .0001); and Initial Highest Count (χ2(1) = 33.41, *p* < .0001). The base for the large model included Initial Highest Count, which resulted in the lowest AIC in our individual model comparisons. The addition of Final Highest Count only marginally improved the fit of this model (χ2(1) = 2.77, *p* = .10), while the addition of Resilience did not produce a better fit to the data (χ2(1) = 0.37, *p* = .54).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hindi** | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC | Model 3: IHC | |
| (Intercept) | -1.97\*\*\* | -2.12\*\*\* | -1.91\*\*\* | -1.88\*\*\* | |
| Resilient | — | 1.79\* | — | — | |
| FHC | — | — | 1.25\*\*\* | — | |
| IHC | — | — | — | 1.41\*\*\* | |
| Trial Within IHC | 1.85\*\*\* | 1.85\*\*\* | 1.76\*\*\* | 1.68\*\*\* | |
| Item Magnitude | -0.85\*\*\* | -0.86\*\*\* | -0.89\*\*\* | -0.92\*\*\* | |
| Age | 0.98\*\*\* | 0.93\*\*\* | 0.76\*\* | 0.53\* | |
| AIC | 853.0 | 850.45 | 827.62 | 821.59 | |
| Conditional R2 | 0.701 | 0.701 | 0.713 | 0.714 | |

Table 11. Base and individual productivity model regression models for predicting Next Number Task performance in Hindi. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**2.5. Gujarati.**Our individual model comparisons indicated that all three candidate measures of productivity significantly predicted Next Number performance (Table 12): Resilience (χ2(1) = 10.06, *p* = .002), Final Highest Count (χ2(1) = 36.25, *p* < .0001), and Initial Highest Count (χ2(1) =33.45, *p* < .0001) all significantly improved the fit of the model in comparison to the base. The large model was constructed using Final Highest Count, which was associated with the lowest AIC in our individual model comparisons, as the base. The addition of Initial Highest Count to this model did not produce a better fit (χ2(1) = 0.87, *p* = .35). The addition of Resilience, however, did significantly explain additional variance (χ2(1) = 4.12, *p* = .04).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Gujarati** | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC | Model 3: IHC | |
| (Intercept) | -1.54\*\*\* | -1.80\*\*\* | -1.47\*\*\* | -1.46\*\*\* | |
| Resilient | — | 2.09\*\*\* | — | — | |
| FHC | — | — | 1.30\*\*\* | — | |
| IHC | — | — | — | 1.25\*\*\* | |
| Trial Within IHC | 2.23\*\*\* | 2.19\*\*\* | 2.03\*\*\* | 1.98\*\*\* | |
| Item Magnitude | -1.16\*\*\* | -1.17\*\*\* | -1.24\*\*\* | -1.26\*\*\* | |
| Age | 0.48\* | 0.44 | 0.37 | 0.32 | |
| AIC | 795.5 | 787.41 | 761.23 | 764.02 | |
| Conditional R2 | 0.695 | 0.696 | 0.710 | 0.710 | |

Table 12. Base and individual productivity model regression models for predicting Next Number Task performance in Gujarati. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**2.6. English (India).**Our individual model comparisons indicated that Final Highest Count (χ2(1) = 21.1, *p* < .0001) and Initial Highest Count (χ2(1) = 19.49, *p* = .0001) significantly predicted Next Number performance (Table 13). We built our large model with Final Highest Count included in the base, as this measure was associated with the lowest AIC in our individual model comparisons. The addition of Initial Highest Count significantly improved the fit of this model (χ2(1) = 5.88, *p* = .02).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **English (India)** | ***Coefficient estimates (β)*** | | | |
| *Parameters* | Base | Model 1: Resilience | Model 2: FHC | Model 3: IHC | |
| (Intercept) | -0.87\*\*\* | -1.10\*\*\* | -0.80\*\*\* | -0.77\*\*\* | |
| Resilient | — | 0.56 | — | — | |
| FHC | — | — | 1.27\*\*\* | — | |
| IHC | — | — | — | 1.06\*\*\* | |
| Trial Within IHC | 1.40\*\*\* | 1.41\*\*\* | 1.23\*\*\* | 1.13\*\*\* | |
| Item Magnitude | -1.25\*\*\* | -1.25\*\*\* | -1.31\*\*\* | -1.37\*\*\* | |
| Age | 1.21\*\*\* | 1.06\*\*\* | 0.25 | 0.57\* | |
| AIC | 697.8 | 698.32 | 678.68 | 680.30 | |
| Conditional R2 | 0.676 | 0.676 | 0.678 | 0.685 | |

Table 13. Base and individual productivity model regression models for predicting Next Number Task performance in Indian English. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

**2.7. Cross-linguistic models*.*** In Experiment 1 (Table 14, Cantonese, Slovenian, and US English), our individual model comparisons indicated that both Resilience (χ2(1) = 19.4, *p* < .0001) and Final Highest Count (χ2(1) = 38.96, *p* < .0001) significantly improved the fit of the base model. The addition of Resilience to a model containing Final Highest Count did not explain additional variance (χ2(1) = 0.27, *p* = .60).

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Comparison to Cantonese** | | | | | **Comparison to Slovenian** | | | | | |
|  | **Coefficient Estimates (*β)*** | | | | | **Coefficient Estimates (*β)*** | | | | | |
| *Predictors* | Base | | Model 1: Resilience | | Model 2: FHC | | Base | | Model 1: Resilience | Model 2: FHC | |
| (Intercept) | -1.38\*\*\* | | -1.80\*\*\* | | -1.51\*\*\* | | 0.21 | | -0.14 | 0.27 | |
| Resilient | - | | 0.88\*\*\* | | - | | - | | 0.88\*\*\* | - | |
| FHC | - | | - | | 1.05\*\*\* | | - | | - | 1.05\*\*\* | |
| Cantonese | - | | - | | - | | -1.58\*\*\* | | -1.66\*\*\* | -1.77\*\*\* | |
| Slovenian | 1.58\*\*\* | | 1.66\*\*\* | | 1.77\*\*\* | | - | | - | - | |
| English (US) | 1.77\*\*\* | | 1.78\*\*\* | | 1.82\*\*\* | | 0.18 | | 0.12 | 0.04 | |
| IHC | 1.27\*\*\* | | 1.26\*\*\* | | 0.65\*\* | | 1.93\*\*\* | | 1.65\*\*\* | 0.88\*\* | |
| Trial Within IHC | 1.20\*\*\* | | 1.19\*\*\* | | 1.18\*\*\* | | 1.20\*\*\* | | 1.19\*\*\* | 1.18\*\*\* | |
| Item Magnitude | -0.91\*\*\* | | -0.92\*\*\* | | -0.92\*\*\* | | -0.91\*\*\* | | -0.92\*\*\* | -0.92\*\*\* | |
| Age | 0.92\*\*\* | | 0.75\*\*\* | | 0.61\*\*\* | | 0.92\*\*\* | | 0.75\*\*\* | 0.61\*\*\* | |
| WPPSI | 0.19\* | | 0.17 | | 0.15 | | 0.19\* | | 0.17 | 0.15 | |
| Cantonese: IHC | - | | - | | - | | -0.65\* | | -0.39 | -0.23 | |
| Slovenian: IHC | 0.65\* | | 0.39 | | 0.23 | | - | | - | - | |
| English (US): IHC | 0.26 | | 0.17 | | 0.13 | | -0.40 | | -0.21 | -0.11 | |
| AIC | 2986.06 | 2968.2 | | 2949.1 | | |  |  | | |  |
| Conditional R2 | 0.708 | 0.706 | | 0.702 | | |  |  | | |  |

Table 14. Base and individual productivity model regression models for predicting Next Number performance in cross-linguistic analyses with Cantonese, Slovenian, and US English, with Cantonese (left) and Slovenian (right) selected as reference groups. *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

In Experiment 2 (Table 15, Hindi, Gujarati, and US English), we again found that both Resilience (χ2(1) = 9.91, *p* = .002) and Final Highest Count (χ2(1) = 22.68, *p* < .0001) significantly improved the fit of the base model. The addition of Resilience to a model containing Final Highest Count did not significantly improve its fit (χ2(1) = .64, *p* = .43). In a *post hoc* analysis with Indian English substituted for US English, we also found that only Final Highest Count significantly improved the fit of the base model (χ2(1) = 8.36, *p* = .004).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Comparison to English (US)** | | | **Comparison to English (India)** | | |
|  | **Coefficient Estimates (*β)*** | | | **Coefficient Estimates (*β)*** | | |
| *Predictors* | Base | Model 1: Resilience | Model 2: FHC | Base | Model 1: Resilience | Model 2: FHC |
| (Intercept) | -0.33 | -0.74\*\* | -0.63\*\* | -0.97\*\*\* | -1.19\*\*\* | -1.33\*\*\* |
| Resilient | - | 0.98\*\* | - | - | 0.50 | - |
| FHC | - | - | 0.98\*\*\* | - | - | 0.66\*\* |
| Hindi | -1.00\*\* | -0.67 | -0.53 | -0.17 | -0.01 | 0.30 |
| Gujarati | -0.93\*\* | -0.66 | -0.53 | -0.14 | 0 | 0.29 |
| IHC | 1.11\*\*\* | 1.01\*\*\* | 0.42\* | 0.63\*\*\* | 0.62\*\* | 0.25 |
| Trial Within IHC | 1.73\*\*\* | 1.73\*\*\* | 1.73\*\*\* | 1.66\*\*\* | 1.67\*\*\* | 1.66\*\*\* |
| Item Magnitude | -0.90\*\*\* | -0.90\*\*\* | -0.91\*\*\* | -1.21\*\*\* | -1.21\*\*\* | -1.21\*\*\* |
| Age | 0.97\*\*\* | 0.85\*\*\* | 0.80\*\*\* | 0.71\*\* | 0.64\*\* | 0.57\*\* |
| WPPSI | 0.20 | 0.20 | 0.15 | 0.20 | 0.18 | 0.13 |
| Hindi: IHC | 1.68\*\*\* | 1.61\*\* | 1.48\*\* | 2.19\*\*\* | 2.13\*\*\* | 2.05\*\*\* |
| Gujarati: IHC | 0.75 | 0.50 | 0.45 | 1.22\*\* | 1.07\*\* | 1.00\*\* |
| AIC | 2558.2 | 2550.3 | 2544.2 | 2146.6 | 2146.5 | 2140.2 |
| Conditional R2 | 0.738 | 0.736 | 0.733 | 0.717 | 0.716 | 0.716 |

Table 15. Base and individual productivity model regression models for predicting Next Number performance in cross-linguistic analyses with Hindi, Gujarati, US English (left) and Indian English (right). *\*p* < .05; *\*\*p* < .01; \*\*\**p* < .001.

1. **Analyses of items above and below 100**

For all languages we tested, 100 marks the introduction of a new decade label. Children must also learn that the introduction of this new decade label must be appended to the beginning of the count list, and does not replace the other decade labels that came prior. Further, most counting instruction likely does not exceed 100, such that children often have to discover these rules on their own. For many children, these syntactic and morphological changes proved to be quite challenging; in many languages children were able to use prompts to count up to 100, but not beyond. Although we observed a decrease in performance on both the Unit and Next Number tasks with increasing item magnitude in every language (Figure 1), we wished to test whether items above 100 as a group were particularly difficult for children in these tasks. In particular, we explored whether an item being less or greater than 100 accounted for unique variance in children’s performance beyond their Initial Highest Count and productive counting knowledge.

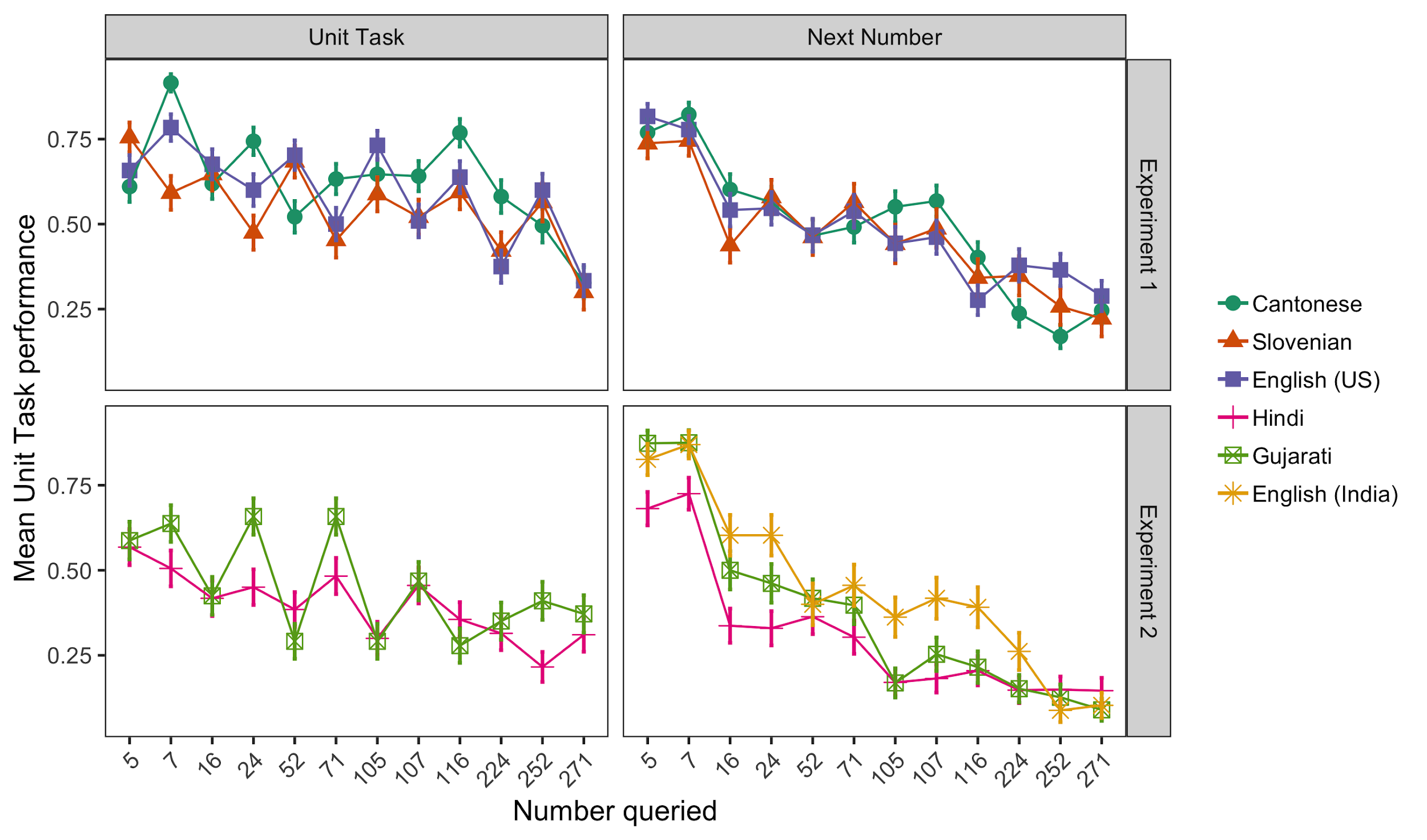


Figure 1. Mean performance by item for the Unit and Next Number tasks in Experiments 1 and 2. Error bars represent standard error of the mean.

To do this, we constructed a base model for both the Unit and Next Number tasks in each language which included the strongest productivity predictors for that task. Because item magnitude and whether the item is above or below 100 overlap, we removed the item magnitude term from our base model. Thus, for each language the generalized linear mixed effects models was: Correct ~ [Productivity predictors] + Within IHC + Age + (1|Subject). We then added a term indicating whether the queried item was above or below 100, and conducted a Likelihood Ratio Test to determine whether this term significantly improved the fit of the model.

**3.1. Unit Task.**The addition of a term indicating whether a queried item was above or below 100 did not significantly improve the fit of the base model in Cantonese (χ2(1) = .0008, *p* = .98), or in Slovenian (χ2(1) = 2.94, *p* = .09). In these two languages, children’s performance was more significantly predicted by whether the item was within or outside their Initial Highest Count. However, adding this term did significantly improve the fit of the model in US English (χ2(1) = 6.50, *p* = .01), Hindi (χ2(1) = 13.35, *p* = .0003), and Gujarati (χ2(1) = 11.89, *p* = .0006), such that performance for items above 100 was significantly worse in each language (US English: *β = -*0.40, *p* = .008; Hindi: *β =* -0.62, *p* = .0003; and Gujarati: *β = -*0.59, *p* = .0006).

**3.2. Next Number.**As in our Unit Task analysis, the addition of a term indicating whether an item was above or below 100 did not significantly improve the fit of the base model in Cantonese (χ2(1) = .12, *p* = .73); once again, children’s performance was better predicted by whether the item was outside their Initial Highest Count than whether the item was above 100. On the other hand, adding this term did significantly improve the fit of the base model in Slovenian (χ2(1) = 45.76, *p* < .0001), US English (χ2(1) = 36.91, *p* < .0001), Hindi (χ2(1) = 42.23 *p* <.0001), and Gujarati (χ2(1) = 60.58, *p* < .0001), such that performance for items above 100 was significantly worse in each language (Slovenian: *β =-1.54, p* < .0001; US English: *β =* -1.31, *p* < .0001; Hindi: *β =* -1.72, *p* < .0001; and Gujarati: *β = -*1.94, *p* < .0001).

Overall, we found some evidence that items above 100 were more difficult for children, particularly in the Next Number task, where children must generate the next number without alternatives. Interestingly, we did not find that items above 100 were more difficult for Cantonese-speaking children, but rather that their performance on both the Unit and Next Number tasks was more closely related to whether an item was within their Initial Highest Count (Unit Task: *β =* .76, *p* < .0001; Next Number: *β =* 2.23, *p* < .0001). This finding likely reflects that a child’s Initial Highest Count was the strongest measure of productivity in Cantonese for both these tasks.

1. **Mean Unit and Next Number Task performance by Resilience classification**

We tested whether our binary Resilience classification broadly captured differences in children’s numerical knowledge by testing mean performance on both the Unit and Next Number tasks within each language. Although we found that graded measures of productive counting knowledge were more strongly predictive of performance on both these tasks, our binary classification nevertheless was individually predictive in most languages using independent samples *t-*tests.

**4.1. Unit Task.**Resilient counters had significantly higher mean Unit Task performance than Non-Resilient counters in: Slovenian (*t*(97) = 5.68, *p* < .0001); US English (*t*(109) = 3.87, *p* = .0002). The difference in mean performance was marginal in Cantonese (*t*(116) = 1.87, *p* = .06) and Hindi (*t*(89) = 1.99, *p* = .05). There was no difference in mean Unit Task performance between Resilient and Non-Resilient counters in Gujarati (*t*(78) = 1.16, *p* = .25).

**4.2. Next Number Task.**Supporting the hypothesis that Resilient counters have some mastery of the productive rules underlying number word generation, we found that Resilient counters had significantly higher mean Next Number performance than Non-Resilient counters in all languages: Cantonese (*t*(116) = 3.22, *p* = .002); Slovenian (*t*(97) = 9.21, *p* < .0001); US English (*t*(107) = 7.71, *p* < .0001); Hindi (*t*(89) = 2.27, *p* = .03); Gujarati (*t*(78) = 3.37, *p* = .001); and Indian English (*t*(68) = 3.68, *p* = .0005).

1. **Highest Contiguous Next Number**

We used Highest Contiguous Next Number, the highest number for which a child could successfully generate the next number in response to a prompt provided that all the previous numbers were correct, as a measure of their knowledge of productive counting rules (Table 16). The lowest Highest Contiguous Next Number possible was 0, meaning that a child made an error on the training trial with 1 (*n* Cantonese = 8; *n* Slovenian = 9; *n* US English = 8; *n* Hindi = 4; *n* Indian English = 5). The highest number possible was 271; this number would indicate perfect performance on this task (*n* Cantonese = 2; *n* Slovenian = 5; *n* US English = 14; *n* Hindi = 6; *n* Gujarati = 2)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Cantonese** | **Slovenian** | **English (US)** | **Hindi** | **Gujarati** | **English (India)** |
| **Overall**  Mean (*SD*)  Median | 55 (74.39)  16 | 50 (82.56)  7 | 66 (98.41)  7 | 30 (70.86)  7 | 34 (61.55)  7 | 43 (65.89)  12 |
| **Resilient**  Mean (*SD*)  Median | 74 (85.31)  24 | 121 (98.20)  107 | 118 (112.61)  62 | 83 (113.59)  30 | 90 (105.40)  52 | 65 (83.72)  24 |
| **Non-Resilient**  Mean (*SD*)  Median | 35 (54.42)  7 | 25 (59.28)  7 | 33 (71.52)  7 | 25 (64.16)  7 | 25 (46.63)  7 | 28 (46.00)  7 |

Table 16. Highest Contiguous Next Number by language and Resilience. Means and medians are rounded.

Children identified as Resilient counters had significantly higher Highest Contiguous Next Numbers in all languages in comparison to Non-Resilient counters, indicating that these children have acquired a productive rule for generating number words (Cantonese (*t*(116) = 2.94, *p* = .004); Slovenian (*t*(97) = 5.84, *p* < .0001); US English (*t*(107) = 4.85, *p* < .0001); Hindi (*t*(89) = 2.23, *p* = .03); Gujarati (*t*(78) = 3.49, *p* = .0008); and Indian English (*t*(68) = 2.41, *p* = .02)).

Next, we tested whether children learning a more transparent count list may be able to generate higher Next Numbers even without demonstrating evidence of having acquired productive counting rules. We tested this in both Experiment 1 (Cantonese, Slovenian, and English) and Experiment 2 (Hindi, Gujarati, and US English) by building a linear regression predicting Highest Contiguous Next Number in Non-Resilient counters. As in our other cross-linguistic analyses, we attempted to control for between-group differences by including both an age and working memory term. The formula for this model was: Highest Contiguous Next Number ~ Language + Age + Working Memory score.

Reflecting the results of our other cross-linguistic analyses with Cantonese, Slovenian, and US English, we did not find a difference in Highest Contiguous Next Number for Non-Resilient counters between Cantonese and Slovenian (*β =* -14.17, *p* = .17) or between Cantonese and English (*β =* 12.79, *p* = .23). Non-Resilient English-speaking children, on the other hand, had significantly higher Highest Contiguous Next Numbers in comparison to Slovenian children (*β =* 26.96, *p* = .01). In Experiment 2 we found that, in comparison to US English, Non-Resilient counters had significantly lower Highest Contiguous Next Numbers in Gujarati (*β* = -34.64, *p* = .008) and Hindi (*β* = -32.32, *p* = .01). We replicated these results using our Indian English dataset as a reference group (Gujarati: *β* = -29.83, *p =* .01; Hindi: *β* = -27.52, *p* = .02). Thus, while we found that training, rather than transparency, predicted differences in Non-Resilient counters’ performance on this task in three relatively transparent languages (Cantonese, Slovenian, and English), we did find that children learning a more opaque count list (Hindi, Gujarati) were less likely to be able to correctly generate the next number in a sequence without an understanding of their count list’s generative syntax.

1. **Initial Highest Count**

A child’s Initial Highest Count is reflective of both the regularity of a count list, as well as frequency of counting training within that language. Thus, we should predict that languages with more regular count lists, and higher levels of count routine exposure (such as Cantonese) should be associated with a greater frequency of high Initial Highest Counts. To test this prediction, we used Gaussian Mixture Modeling to identify and quantify Initial Highest Count distributions in each language (Figure 2). These models were fit using the ‘mclust5’ package in R (Scrucca, Fop, Murphy, & Raftery, 2017), and aggregated over both Resilient and Non-Resilient Highest Counts. Gaussian Mixture Models (GMM) identify clusters within data (assuming an underlying Gaussian distribution), calculate their mean and variance, and estimate the likelihood of a given point being contained within a given cluster. The number of optimal clusters is selected on the basis of Bayesian Information Criterion. The fits of individual clusters are shown in Table 17.

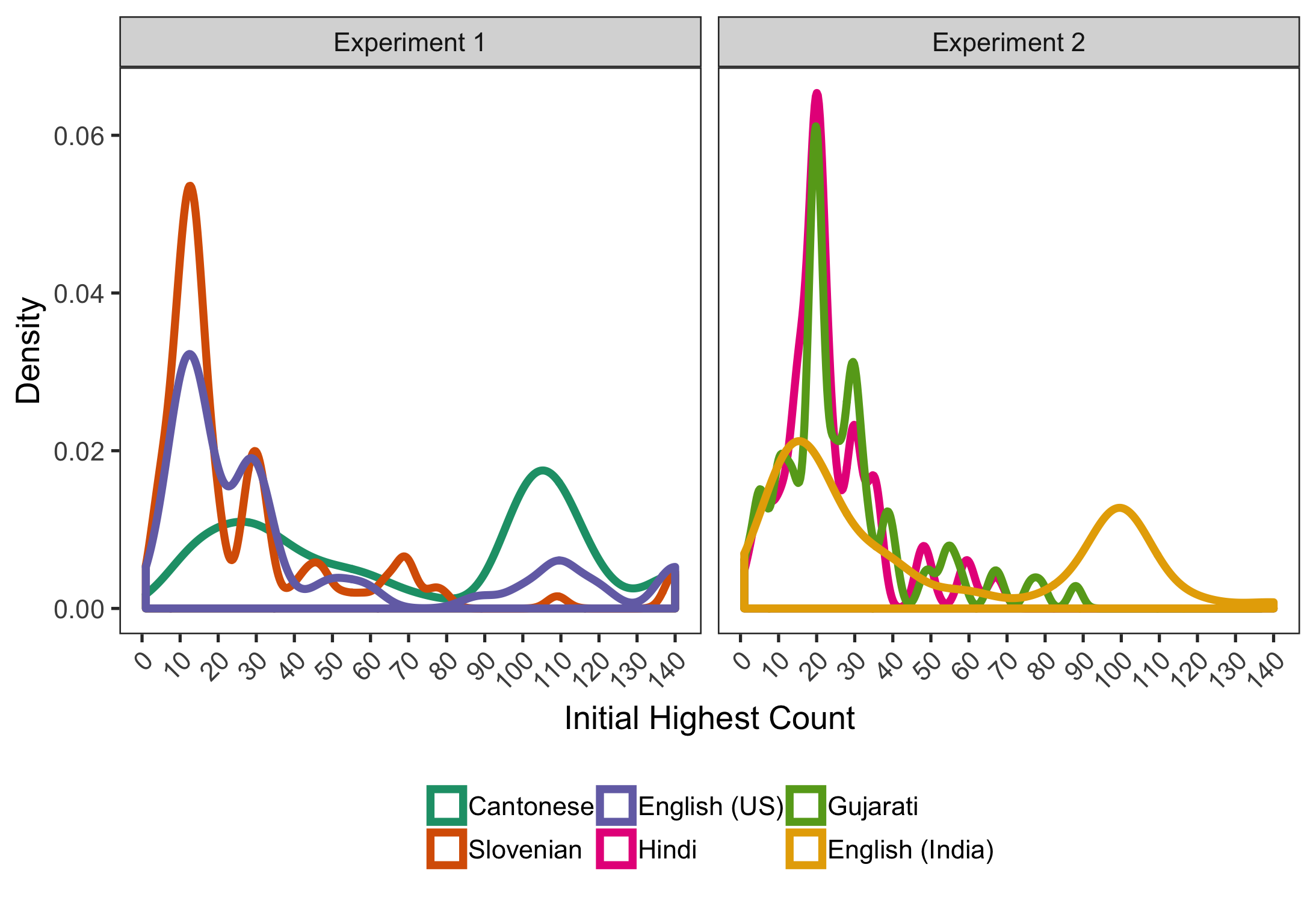


Figure 2. Distribution of Initial Highest Counts in Experiments 1 and 2.

The results of the GMM support our hypothesis that greater levels of training and count list transparency should result in a greater frequency of higher Initial Highest Counts. Cantonese counters’ most frequent Initial Highest Count was 106 (Cluster 4), with a number of much lower probability clusters. In languages with either lower levels of exposure (Slovenian) or transparency (English) the most probable clusters were in the teens (Cluster 1), indicating the challenge this decade poses in extracting recursive counting rules. Indian English clusters were similar to US English clusters in both means and probabilities. In both Hindi and Gujarati, cluster membership largely overlapped with our Resilient/Non-Resilient classification.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Cluster 1** | **Cluster 2** | **Cluster 3** | **Cluster 4** | **Cluster 5** |
| **Cantonese (*n*)**  Mean (*SD*)  Probability  Proportion Resilient | *n* = 20  17.83 (6.57)  0.192  .20 | *n* = 24  34.18 (6.57)  0.189  .46 | *n* =15  59.23 (6.57)  0.119  .73 | *n* = 50  106.12 (6.57)  0.423  .52 | *n* = 20  139.89 (6.57)  0.077  1.0 |
| **Slovenian (*n*)**  Mean (*SD*)  Probability  Proportion Resilient | *n* = 48  11.64 (4.52)  0.49  .08 | *n* = 15  13.58 (1.00)  0.13  0 | *n* = 16  29.73 (1.73)  0.15  .44 | *n* = 20  66.46 (35.05)  0.23  .75 | — |
| **English (US) (*n*)**  Mean (*SD*)  Probability  Proportion Resilient | *n* = 54  12.55 (3.75)  0.47  .13 | *n* = 22  28.91 (0.40)  0.19  .50 | *n* = 35  89.49 (39.24)  0.33  .69 | — | — |
| **English (India) (*n*)**  Mean (*SD*)  Probability  Proportion Resilient | *n* = 32  14.45 (4.94)  0.40  .22 | *n* = 15  33.42 (15.16)  0.27  .67 | *n* = 23  101.01 (11.60)  0.33  .48 | — | — |
| **Gujarati (*n*)**  Mean (*SD*)  Probability  Proportion Resilient | *n* = 69  20.48 (9.22)  0.82  .06 | *n* = 11  56.50 (17.34)  0.18  .64 | — | — | — |
| **Hindi (*n*)**  Mean (*SD*)  Probability  Proportion Resilient | *n* = 82  20.11 (8.40)  0.90  .05 | *n* = 9  55.81 (8.20)  0.10  .44 | — | — | — |

Table 17. Clusters for IHC in each language. Each cluster identifies a mode within the IHC distribution for each language. Mean and SD refer to the mean and SD of that cluster, while probability indicates the likelihood of a given point within that dataset falling into that cluster. Proportion Resilient indicates the proportion of counters in that cluster who met the criteria for Resilience.

Next, we tested whether children need to count higher in languages with less transparent count lists in order to acquire a recursive rule (Yang, 2016) by constructing a linear regression predicting Initial Highest Counts for Non-Resilient counters by language, controlling for age and working memory. Contra this prediction, in Experiment 1 we found that Cantonese Non-Resilient counters had significantly higher IHCs than Slovenian (*β =* 46.26, *p* < .0001) and English Non-Resilient counters (*β =* 20.51, *p* < .0001). We found again that Non-Resilient counters in more transparent languages were able to count significantly higher before acquiring a productive rule in Experiment 2. US English Non-Resilient counters’ had significantly higher Initial Highest Counts in comparison to both Hindi (*β =* 19.04, *p* < .0001) and Gujarati (*β =* 19.40, *p* < .0001). We replicated this finding in a within-culture comparison: Indian English Non-Resilient counters had significantly higher Initial Highest Counts than both Hindi (*β =* 27.79, *p* < .0001) and Gujarati (*β =* 27.57, *p* < .0001) Non-Resilient counters. While the significantly lower Initial Highest Counts of Hindi and Gujarati indicate that count list morphology plays a significant role in extractive recursive counting rules, these findings suggests that this is not purely the result of grammatical structure, as Cantonese-speaking children were able to count quite high prior to demonstrating productive counting knowledge.

1. **Counting errors**

Previous work (Fuson, 1988; Miller & Stigler, 1987; Siegler & Robinson, 1982) has found that children’s errors in the count routine are nonrandom, and that many children are more likely to make errors on decade transitions, which require remembering a new decade label. This pattern of errors would suggest that these children have mastered the decade structure, but have not yet committed the decade labels to memory. We explored whether Resilient Counters differ from Non-Resilient counters in the pattern of their counting errors, hypothesizing that children who are productive are more likely to make errors which involve the recall of a new decade label rather than errors mid-decade, which would indicate that they have not yet mastered the base-10 system.

We found that Resilient Counters made more frequent use of more prompts than Non-Resilient Counters, but the kinds of errors made by both counters varied across languages (Figure 3). In Cantonese and English, a higher proportion of Resilient counters’ errors (about 70%) were at decade transitions (e.g., 49 to 50) than either decade-beginning (Cantonese: 12%; English; 13%) or mid-decade (Cantonese: 17%; English: 17%). Resilient counters in Slovenian, however, had their highest proportion of errors for mid-decade numbers (50%), with comparatively fewer errors at decade transitions (35%) or beginnings (15%). Hindi and Gujarati Resilient counters also had more frequent decade-transition errors. In both Hindi and Gujarati, the syntax for a decade transition already incorporates the next decade label. For example, in Hindi the progression from 38-40 would be *adh****tis*** (38)*, untaa****lis*** (39)*, cha****lis*** (40)*,* where *tis* and *lis* mean *thirty* and *forty* respectively. In line with our hypothesis that Resilient Counters are more likely to require prompts when transitioning to a new decade label, we find that the majority (43%) of Gujarati Resilient counters’ prompts were provided at these points, while these errors comprised 30% of Hindi Resilient counters’ stopping points.

Like Slovenian Resilient counters, however, Hindi and Gujarati Resilient counters were also likely to make errors both in the middle and start of decades. In fact, Resilient Hindi Counters made the highest proportion of errors (45%) mid-decade. This high proportion of mid-decade errors is unsurprising, however, given the irregularity within individual decades: For example, in both Hindi and Gujarati, the decade label for *fifty* alternates between *van* and *pan*. Similarly, Slovenian Resilient counters’ relatively higher frequency of mid-decade errors may reflect their lower levels of counting training.

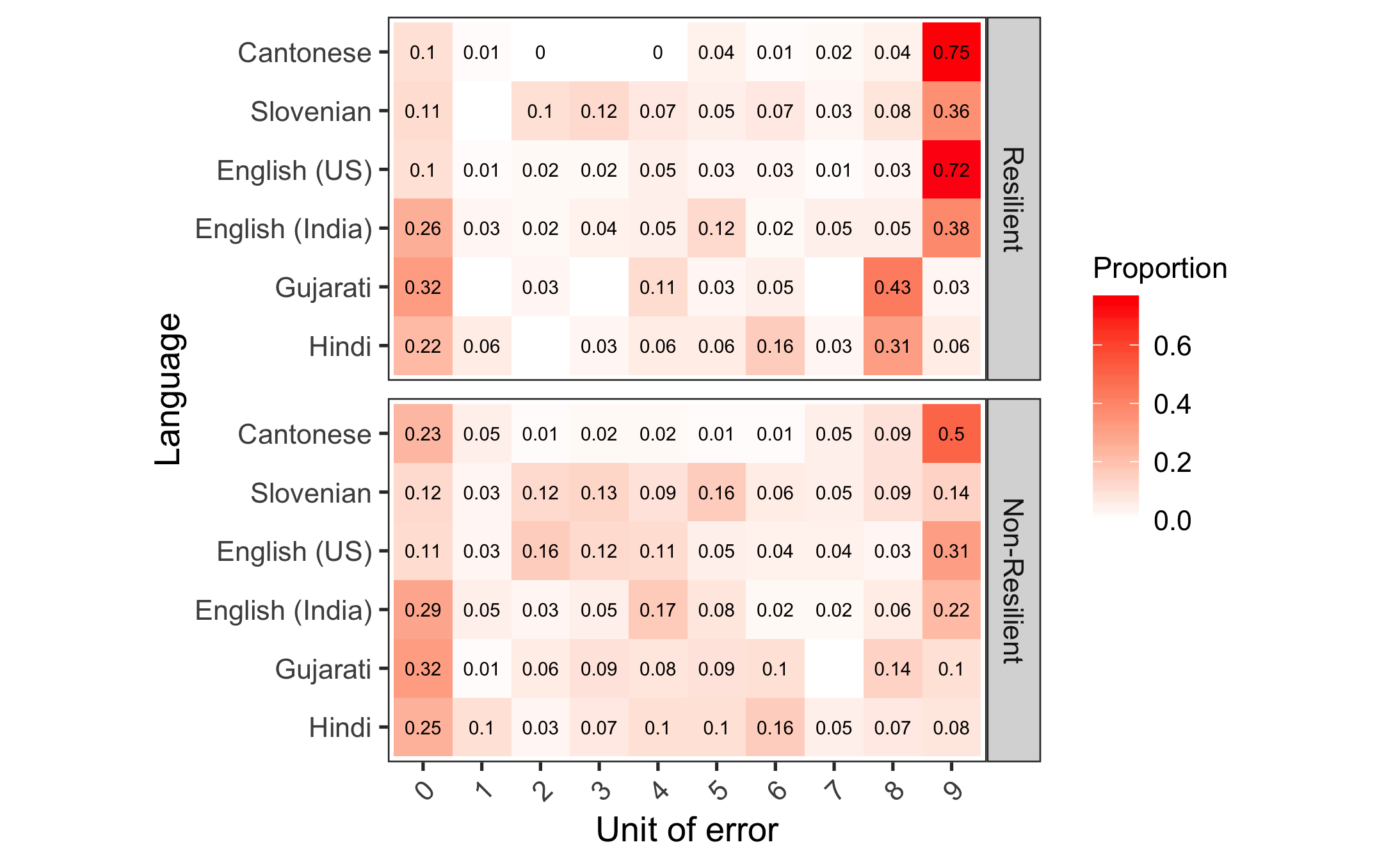


Figure 3. Frequency of unit in which children made an error in each language, grouped by Resilience. Color corresponds to the proportion of total errors made by Resilient or Non-Resilient counters within each language. The unit of error corresponds to a child’s last successful count prior to making an error.