RULEX-EM:

Incorporating exemplars and memory effects in a hypothesis-testing model of category learning

Ronald S. Chong (rchong@gmu.edu)
Humans Factors and Applied Cognition
Department of Psychology
George Mason University

Acknowledgements
Robert E. Wray

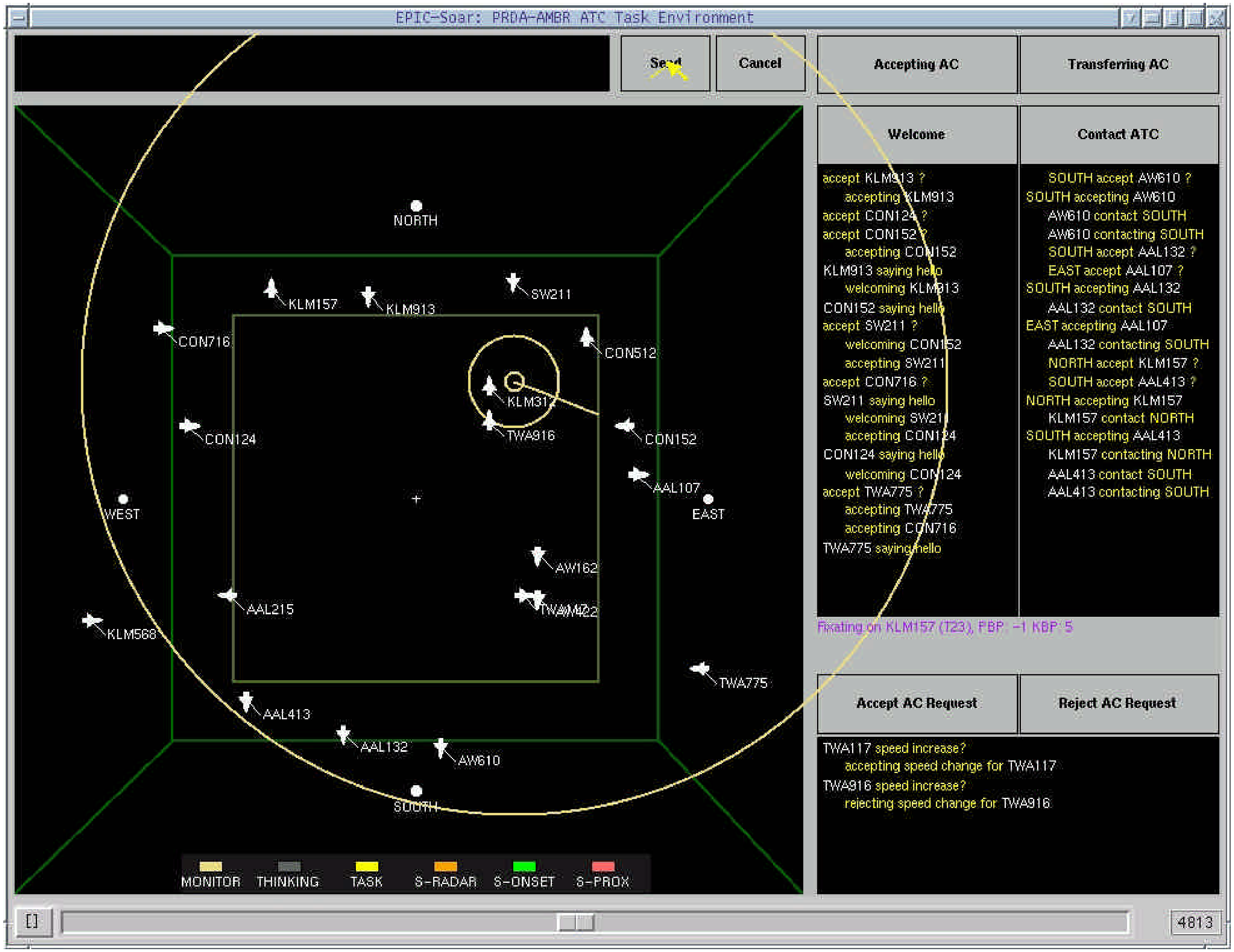
Modeling with EASE

(Elements of ACT-R, Soar, and EPIC)

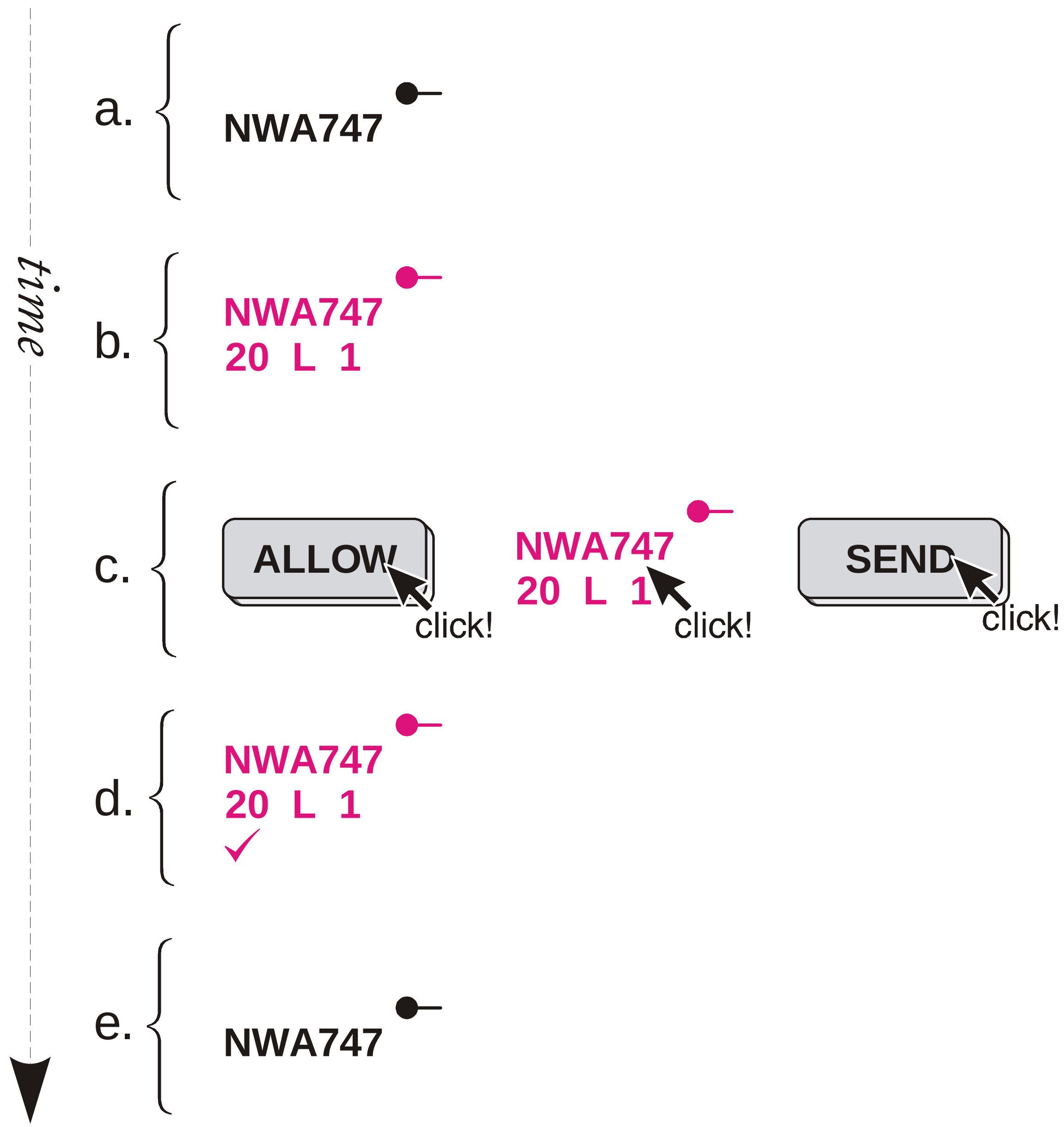
Ronald S. Chong (rchong@gmu.edu)
Humans Factors and Applied Cognition
Department of Psychology
George Mason University

Acknowledgements
Robert E. Wray

DEMO: THE TASK



SCHEMATIC OF A CATEGORY LEARNING TRIAL



- **Category task:**
  - Three features with two values each: FUEL (20 or 40), SIZE (L or S), TURBULENCE (1 or 3)
  - Eight possible instances; 2^3
  - Category is either ALLOW or DENY with four instances in each category
- **Three categorization problems *types*:**
  - **Type 1**: category is defined by a single dimension; e.g. if SIZE is L, then ALLOW. This is the easiest problem
  - **Type 3**: can be characterized as requiring a single-feature rule, plus exception rules.
  - **Type 6**: the most complex category; all features are relevant; correct rules must test three features.

- Type 1

FUEL	SIZE	TURB		CATEGORY
20	S	1	➠	Accept
20	S	3	➠	Accept
20	L	1	➠	Accept
20	L	3	➠	Accept
40	S	1	➠	Reject
40	S	3	➠	Reject
40	L	1	➠	Reject
40	L	3	➠	Reject

- Type 3

FUEL	SIZE	TURB		CATEGORY
20	S	1	➠	Accept
40	S	1	➠	Accept
40	L	1	➠	Accept
20	S	3	➠	Accept
20	L	3	➠	Reject
40	S	3	➠	Reject
40	L	3	➠	Reject
20	L	1	➠	Reject

- Type 6

FUEL	SIZE	TURB		CATEGORY
20	S	3	➠	Accept
20	L	1	➠	Accept
40	S	1	➠	Accept
40	L	3	➠	Accept
20	S	1	➠	Reject
40	S	3	➠	Reject
40	L	1	➠	Reject
20	L	3	➠	Reject

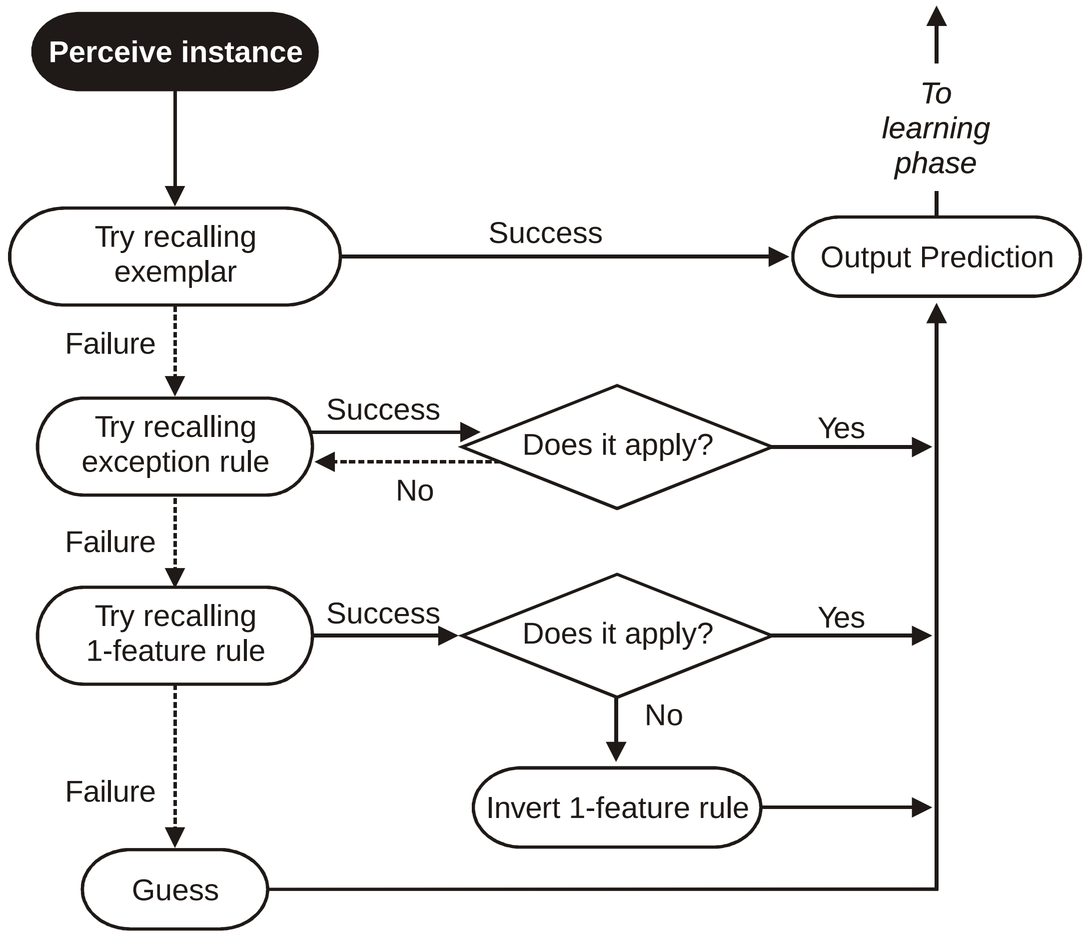
- **Existing models that have been fit to Nosofsky's data:**
  - RULEX, ALCOVE, Configural Cue, Configural Cue w/ DALR, SUSTAIN, rational model
- **Why not use an existing category learning model?**
  - Very few are process models
  - None are implemented in a cognitive architecture
  - Time-to-categorize is not a typical output of these models
  - Lots of free parameters; as many as ten in one model
  - Few hybrid models (containing both exemplars and rules)
  - Few models represent memory effects (forgetting)
  - None are sensitive to time; e.g. inter-stimulus time
  - None are sensitive to the presence of a secondary task
- **Goal: Build a model that does all this**

- A process model
- Implemented in a cognitive architecture (EPIC-Soar)
- Inspired by RULEX, a hypothesis-testing process model
- Incorporates rules and exemplars
- Forgetting, using an ACT-R mechanism
- Uses a smaller set of parameters
- "-EM" means Exemplars and Memory constraints

- **Derived from ACT-R's mechanism**
- **Parameters used for this model:**
  - decay-rate = -0.5
  - transient noise = 0.25
  - retrieval threshold = 0.0
  - base-level constant = 1.0
- ***These are all ACT-R default parameters or commonly used values for ACT-R models.***

- **Like RULEX and other models, this model uses a homogeneous representation for exemplars and rules:**
- **Four-tuple:**
  - One slot for each feature; e.g. fuel, size, turbulence
  - One for the category
- **Two kinds of rules: single-feature and exceptions**
- **Examples:**
  - *Exemplar*: [FUEL = 20; SIZE = S; TURB = 3; CATEGORY = ALLOW]
  - *Single-feature*: [FUEL = *; SIZE = *; TURB = 3; CATEGORY = ALLOW]
  - *Exception*: [FUEL = *; SIZE = L; TURB = 3; CATEGORY = DENY]
- **These structures are all subject to decay and forgetting.**

- Mostly inherited from RULEX
- Strict sequential use of category prediction strategies
- New part is "Try recalling exemplar"



Prediction by guessing on first trial...

```
1: O: O1 (perceive-instance)
1: instance 0: 4000 L 3
2: O: O2 (failed-episodic-recall)
2: unable to recall a classification
3: O: O3 (failed-exception-rules)
3: unable to recall an exception
4: O: O4 (failed-1-dim-rules)
4: unable to recall a 1-dim rule
5: O: O6 (guess-reject)
6: O: O7 (output-prediction)
6: sending prediction: R
```

Generating a prediction by recalling the classification...

```
49420: O: O1019 (perceive-instance)
49420: instance 64: 4000 S 3
49421: O: O1020 (try-episodic-recall)
49421: successfully recalled the classification: R
49422: O: O1022 (output-prediction)
49422: sending prediction: R
```

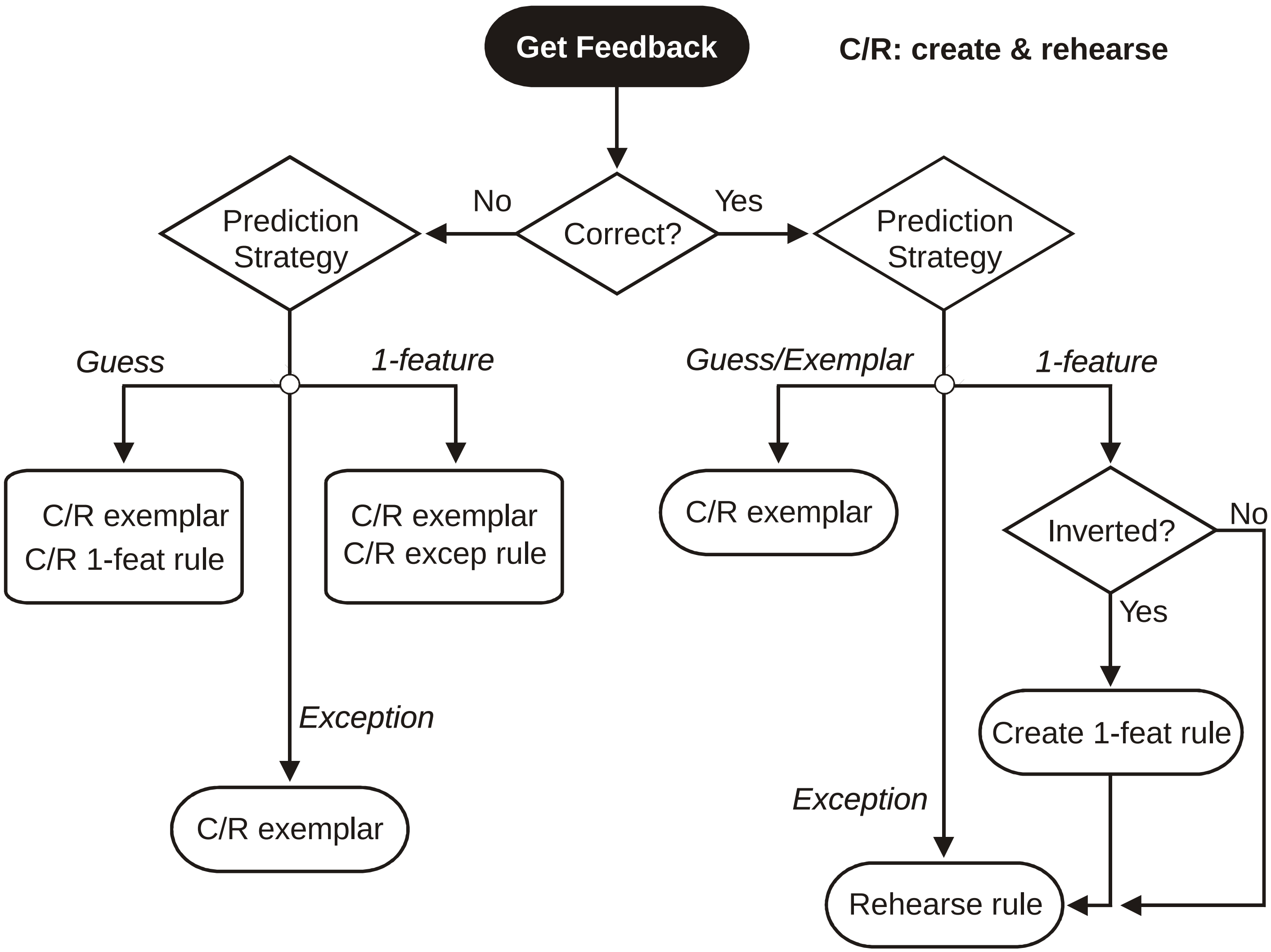
Using a one-dimension rule....

```
7587: O: O136 (perceive-instance)
7587: instance 10: 4000 S 1
7588: O: O137 (failed-episodic-recall)
7588: unable to recall a classification
7589: O: O138 (failed-exception-rules)
7589: unable to recall an exception
7589: available 1-dim rule: (size L ==> A)
7590: O: O139 (try-1-dim-rules)
7590: attending to most active rule: (size L ==> A)
7591: O: O139 (try-1-dim-rules)
7591: it looks like i can use this rule.
7592: O: O139 (try-1-dim-rules)
7592: it cannot be applied directly; will invert
7593: O: O139 (try-1-dim-rules)
7593: will try the inverted form instead: (size S ==> R)
7594: O: O141 (output-prediction)
7594: sending prediction: R
```

Demonstrates strategy of "inverting" a one-dimension rule.

Using an exception rule...

```
14792: O: O275 (perceive-instance)
14792: instance 19: 4000 L 3
14793: O: O276 (failed-episodic-recall)
14793: unable to recall a classification
14793: available exception rule: (size L turbulence 1 ==> R)
14794: O: O277 (try-exception-rules)
14794: attending to most active rule: (size L turbulence 1 ==> R)
14795: O: O277 (try-exception-rules)
14795: oops...rule cannot be applied
14796: O: O277 (try-exception-rules)
14796: available exception rule: (size S turbulence 3 ==> A)
14797: O: O279 (try-exception-rules)
14797: attending to most active rule: (size S turbulence 3 ==> A)
14798: O: O279 (try-exception-rules)
14798: winning exception rule cannot be applied
14799: O: O279 (try-exception-rules)
14800: O: O278 (failed-exception-rules)
14800: available 1-dim rule: (size S ==> R)
14800: available 1-dim rule: (size L ==> A)
14801: O: O280 (try-1-dim-rules)
```


Get Feedback
C/R: create & rehearse
Correct?
No
Yes
Prediction Strategy
Prediction Strategy
Guess
1-feature
Guess/Exemplar
1-feature
C/R exemplar
C/R 1-feat rule
C/R exemplar
C/R excep rule
C/R exemplar
Inverted?
No
Yes
Exception
Create 1-feat rule
C/R exemplar
Exception
Rehearse rule

Memorizing and rehearsing exemplar...

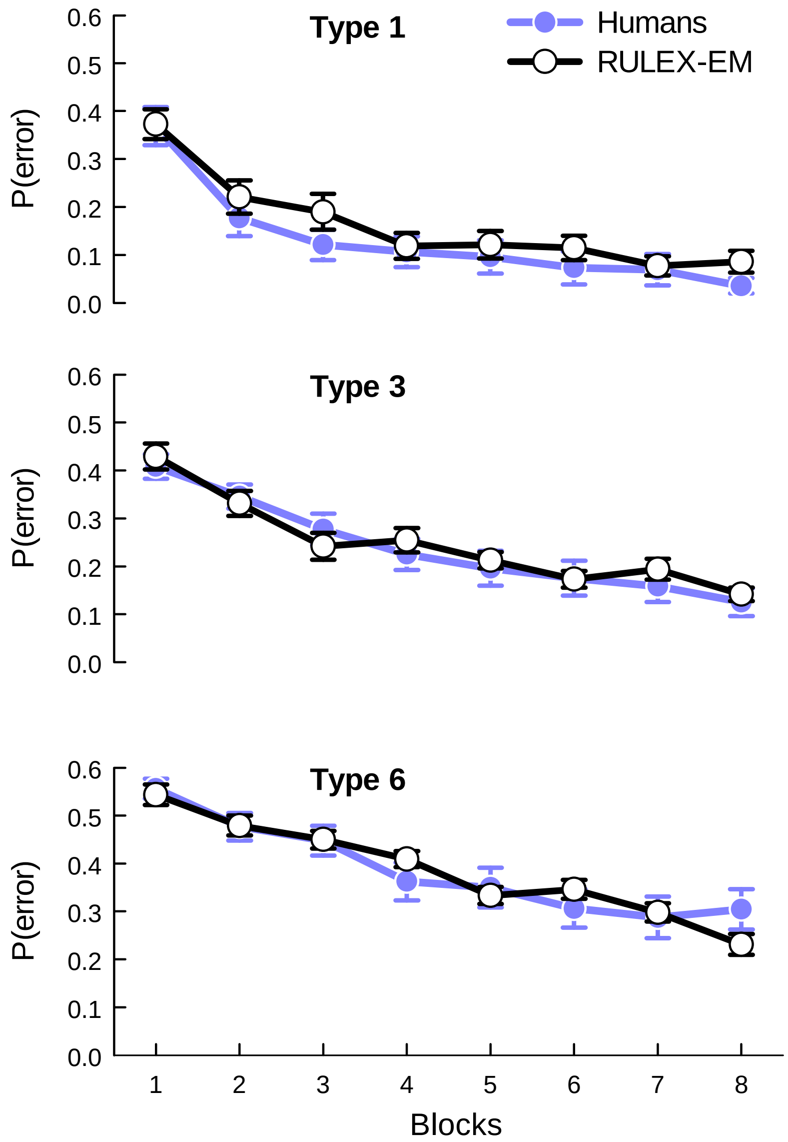
```
1072: O: O23 (guess-reject)
1073: O: O24 (output-prediction)
1073: sending prediction: R
1074: O: O25 (get-feedback)
1074: feedback on trial 1 : CORRECT
1075: O: O26 (acknowledge-correct-prediction)
1076: O: O27 (memorize-classification)
1076: associating correct prediction with the stimulus
1077: O: O28 (rehearse-classification)
1077: rehearsing classification
1082: O: O29 (clean-up)
```

Learning a 1-dim rule...

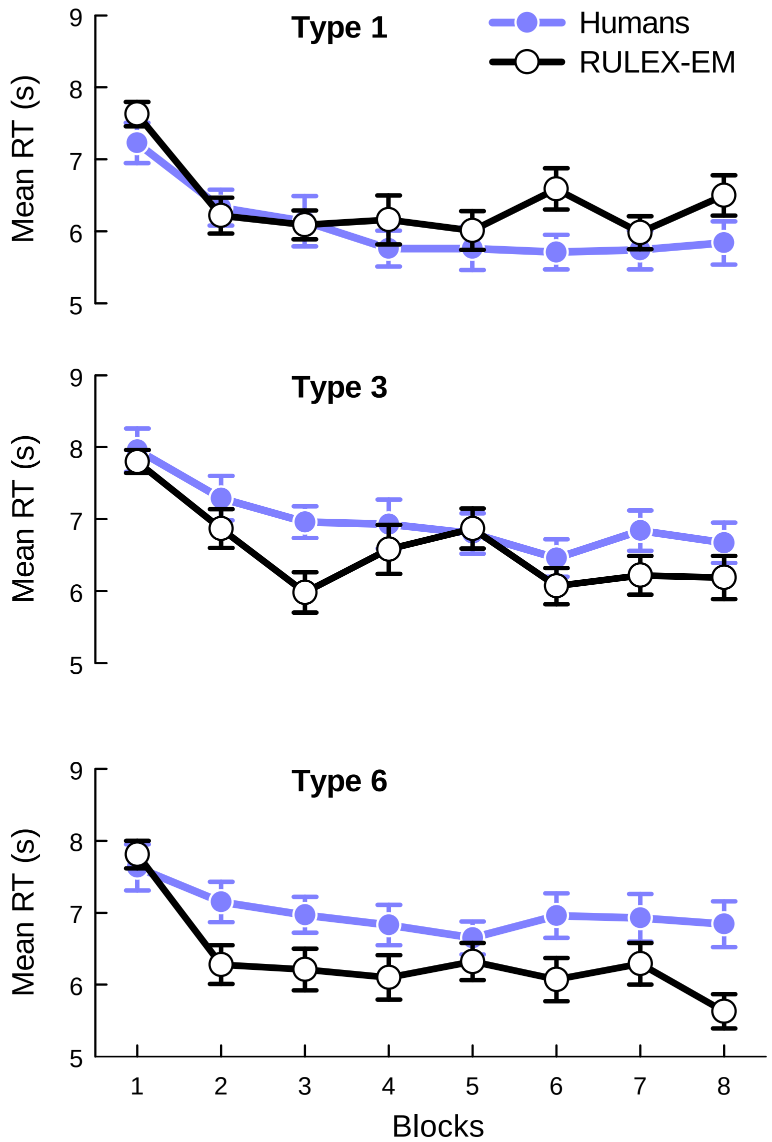
```
5: O: O6 (guess-reject)
6: O: O7 (output-prediction)
6: sending prediction: R
7: O: O8 (get-feedback)
7: feedback on trial 0 : INCORRECT
8: O: O9 (derive-correct-prediction)
9: O: O10 (memorize-classification)
9: associating correct prediction with the stimulus
10: O: O11 (rehearse-classification)
10: rehearsing classification
15: O: O13 (sample-dim-for-1-dim-rule)
16: O: O15 (create-1-dim-rule)
16: building 1-dim rule: elaborating size with L
17: O: O15 (create-1-dim-rule)
18: O: O15 (create-1-dim-rule)
18: memorizing 1-dim rule: (size L ==> A)
19: O: O16 (rehearse-rule)
19: rehearsing rule: (size L ==> A)
27: O: O17 (clean-up)
```

Learning an exception rule...

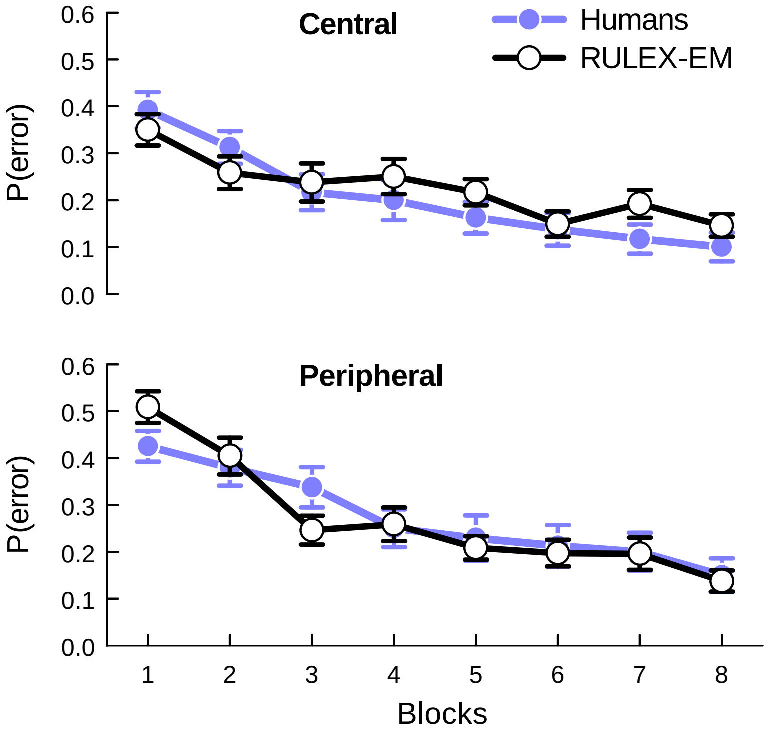
```
17838: attending to most active rule: (size S ==> R)
...
17841: O: O341 (try-1-dim-rules)
17841: successfully applying the attended 1-dim rule
17842: O: O343 (output-prediction)
17842: sending prediction: R
17843: O: O344 (get-feedback)
17843: feedback on trial 23 : INCORRECT
17844: O: O345 (derive-correct-prediction)
17845: O: O346 (note-failed-dim-in-1-dim-rule)
17845: memorizing failed-dim-for-1-dim-rule: size
17846: rehearsing state-info: failed-dim-for-1-dim-rule size
...
17855: O: O349 (sample-1st-dim-for-exception)
17856: O: O350 (sample-other-dims-for-exception)
17857: O: O352 (create-exception)
17857: building exception: size S turbulence 1
17859: O: O352 (create-exception)
17859: memorizing exception rule: (size S 1 ==> A)
17860: O: O353 (rehearse-rule)
17860: rehearsing rule (size S 1 ==> A)
17868: O: O17 (clean-up)
```



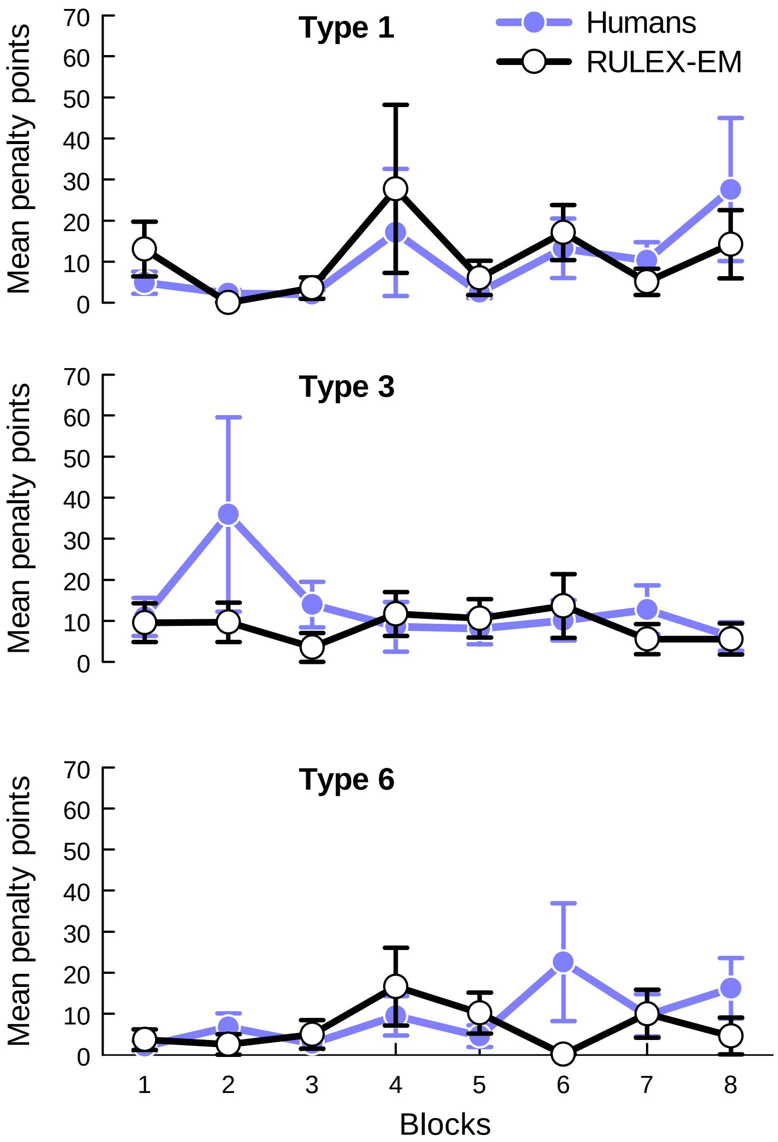
- Satisfactory fit.
- $G^2 = 5.64$



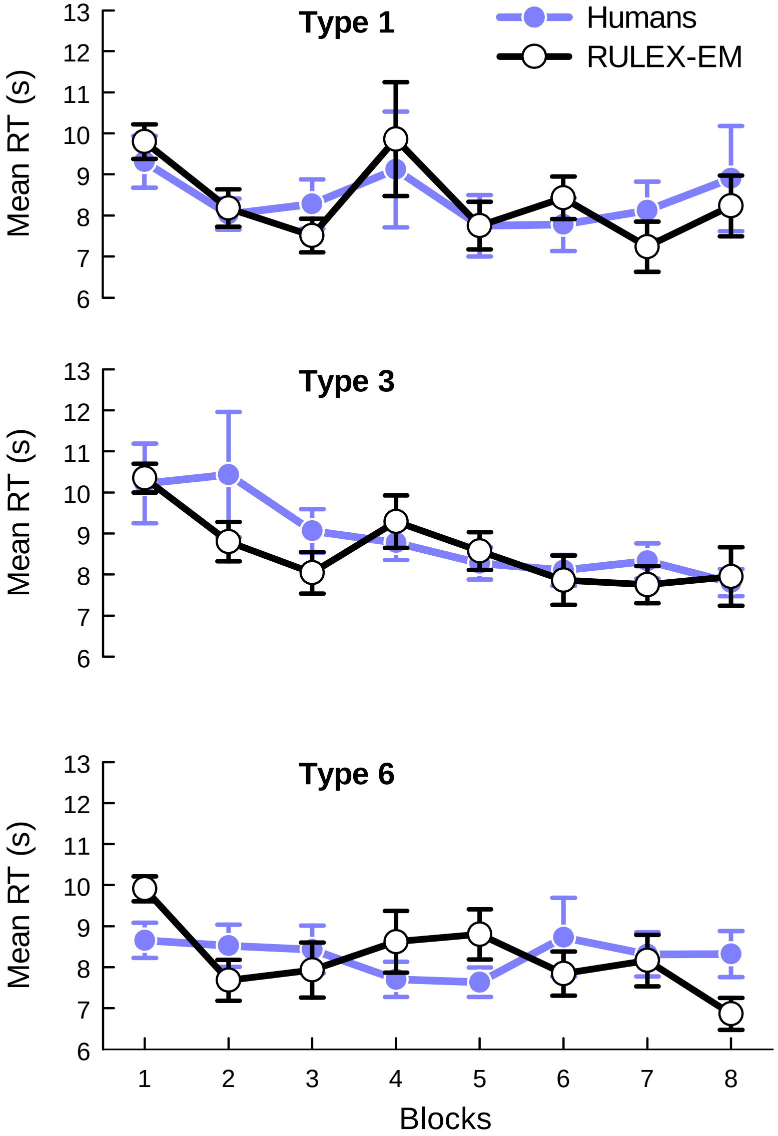
- SSE = 8.40
- Why is the model faster on Category 6?
  - Exemplar and exception-rule recall account for 80% of responses.
  - These are the first two prediction strategies, so prediction ends relatively early.



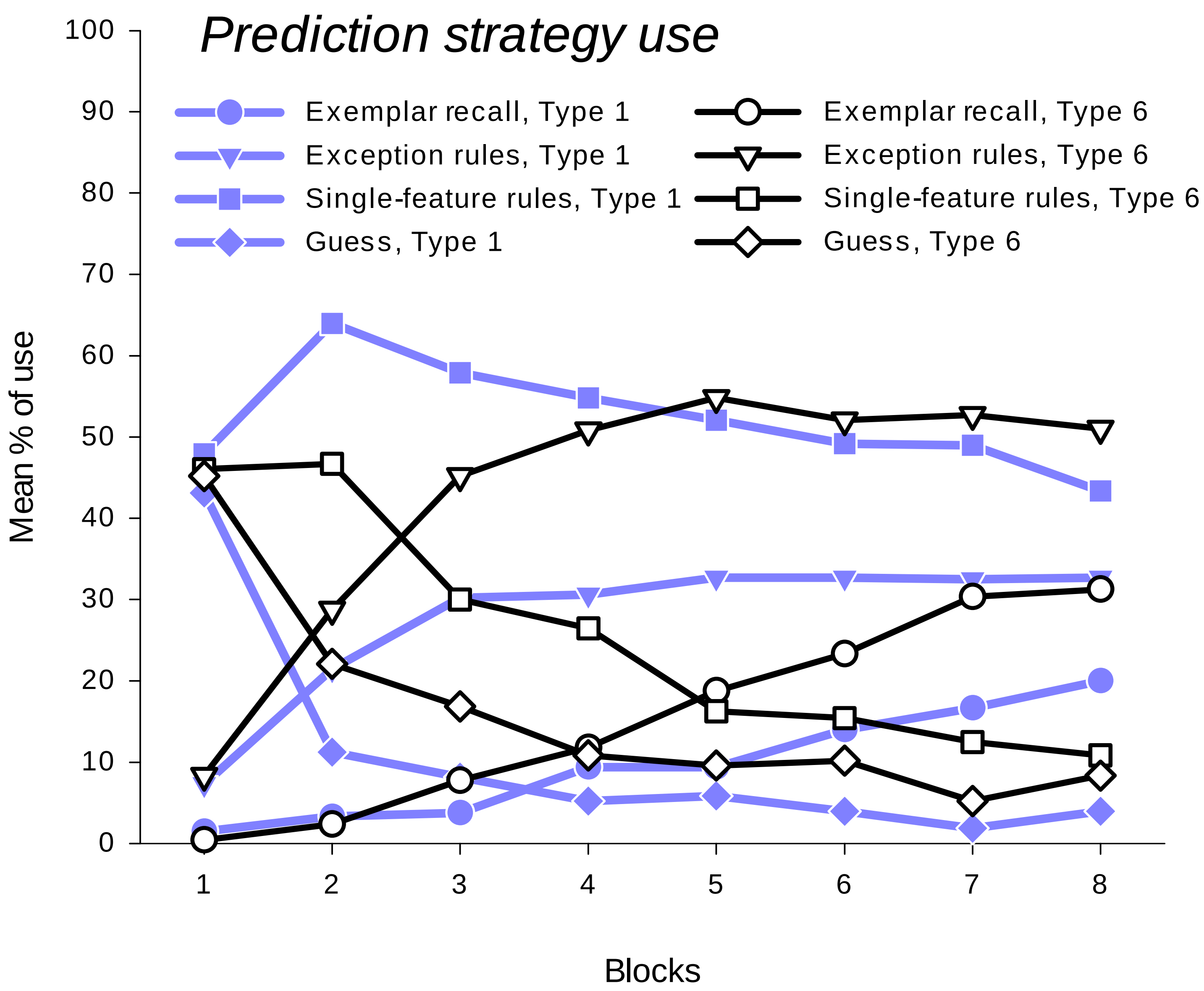
- **Satisfactory fit.**
- $G^2 = 5.89$



- SSE = 2046
- Notice that the model often gives a *qualitative* prediction of performance variability.


Type 1
Humans
RULEX-EM
Mean RT (s)
13
12
11
10
9
8
7
6
Type 3
Type 6
1
2
3
4
5
6
7
8
Blocks

- SSE = 15.24


Prediction strategy use
Exemplar recall, Type 1
Exception rules, Type 1
Single-feature rules, Type 1
Guess, Type 1
Exemplar recall, Type 6
Exception rules, Type 6
Single-feature rules, Type 6
Guess, Type 6
Mean % of use
0
10
20
30
40
50
60
70
80
90
100
1
2
3
4
5
6
7
8
Blocks

- **One validation of EASE.**

  "...incorporating the mechanisms of other architectures and models and 'inheriting' their validation against human data promises to result in rapid progress as parallel developments by other architectures emerge." (Pew & Mavor, 1998, p. 95).

- **Parameters manipulated to achieve fits:**
  - # of rehearsals for memorizing exemplar; final value = 4.
  - # of rehearsals for memorizing rules; final value = 7.

- **Model was *fitted* only to P(error) by problem type; all others were *predictions* of the model.**

- **Tons of empirical category learning data for further validation.**

- Tons of empirical category learning data for further validation.
- Does not capture the different strategies subject make take; i.e. "On this trial, I'm just going to memorize the stimuli."