**Andy’s Talk Gregynog 2024**

**Title: Different Perspectives on the Idea of a “Reward Prediction Error”**

**Abstract:** The reward prediction error (RPE) concept has been influential in both associative learning and reinforcement learning theories, with new learning depending upon such errors. For years, we have assumed that the difference between obtained and expected rewards (e.g., l – SV) drive new learning (although our theories have sometimes expanded this idea to include events other than rewards, as well-imagined, e.g., in MKM). Since the connectionist movement in the 1980s, it was made clear that some mechanism is needed to propagate RPE signals throughout a complex neural network that consists of more than two layers. Since the brain consists of a complex neural network of interacting structures and is the primary organ through which learning takes place, it is time for associative learning theory to acknowledge and address this fundamental problem. One solution, apparently used widely today in AI, was to back-propagate the RPE signal (computed at an output layer) throughout the entire network. While that solution works at an engineering level, it is far from biologically plausible. An alternative framework has been to consider that some neurobiological process, e.g., dopamine, computes the RPE and then broadcasts this signal widely throughout the brain enabling new plasticity at any given synapse that receives the signal. Yet another approach, not widely recognized but is more in the spirit of Konorski (1948), is referred to as “Contrastive Hebbian Learning.” This algorithm plausibly reconceptualizes how RPEs might arise in complex networks by contrasting the activity levels of any two conceptual units (e.g., neurons) before and after reward is presented. This, effectively, means that connection strengths will increase (or decrease) when a sending unit is paired with a rise (or fall) in the activation level of a receiving unit. Thus, each conceptual “synapse” within a complex network has its own locally computable RPE term, but whose individual elements (l and SV) may very well change throughout the course of learning. While it is the goal of the system to decrease error discrepancies, the approach emphasizes the distributed nature of the reward representation and does not so much focus on whether a particular event representation (like reward) is anticipated to match reality. Instead, this algorithm ensures that the entire activity pattern of a complex interacting network produced by reward is, itself, reproduced by presentation of the predicting stimuli. In this sense, the RPE signal is not so much about predicting an “event” as it is the impact that such an event will have on an entire network. This strikes us as an insightful way of conceptualizing the RPE concept. I will review some of our efforts to examine this more biologically plausible learning algorithm applied to multi-layer networks and show that it produces results consistent with many of the associative learning phenomena our field takes to be important.

**Santiago’s Poster Gregynog 2024**

**Title: Neurocomputational plausibility in Categorical Reversal Learning: Contrastive Hebbian Learning with Dynamic learning rates**

**Abstract:** Backpropagation (BP) is one of the most famous and widely used artificial neural network (ANN) algorithms. However, BP has been criticized for its lack of neurobiological plausibility, where the brain cannot compute backward error signals (*e.g.*, Lillicrap, et al, 2020 – Nature Review Neuroscience). There is a distinct family of ANN algorithms based on the Hebbian learning rule (“fire together wire together”), these ANNs can also solve most of learning problems as BP. This family is known as Contrastive Hebbian Learning (CHL) and instead of sending a discrepancy (or prediction error) learning signal backwards, it computes learning signals locally at every neurocomputational unit. Thus, CHL is a more biologically plausible algorithm. Although, neither CHL nor BP have an adequate prediction for category formation, nor have an explanation for it. Category formation can be studied with a Categorical Reversal Learning (CRL) task. In total-reversal condition, living agents learn that half exemplars belong to category A and the other half belong to category B. Then in a second phase, the exemplars from category A change to category B, and vice versa. Agents learn total-reversal categories faster than a comparison partial-reversal condition, where not all the exemplars change from one category the other in phase two. Interestingly, ANN learn faster partial-reversals than total-reversals, opposite as living agents. Delamater (2023 – Gregynog) proposed an improvement to BP that incorporates two ‘opposed’ attentional mechanisms based on adaptive/dynamic learning rates – Pearce & Hall (1980) and Mackintosh (1975). In this work we (1) present data corroborating that humans are better at learning total-reversals than partial-reversals, and (2) we incorporate the adaptive/dynamic learning rates to CHL to create a biologically plausible model that learns categories as humans may do it. Finally, we present our ongoing project ALANN (Associative Learning with Artificial Neural Networks), an open-source initiative to study learning theories with connectionist models: <https://github.com/santiagocdo/ALANN>.

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. Nature Reviews Neuroscience, 21(6), 335-346. <https://www.nature.com/articles/s41583-020-0277-3>

Shervani-Tabar, N., & Rosenbaum, R. (2023). Meta-learning biologically plausible plasticity rules with random feedback pathways. Nature Communications, 14(1), 1805. <https://www.nature.com/articles/s41467-023-37562-1>

Liu, Y. H., Smith, S., Mihalas, S., Shea-Brown, E., & Sümbül, U. (2022). Biologically-plausible backpropagation through arbitrary timespans via local neuromodulators. Advances in Neural Information Processing Systems, 35, 17528-17542. <https://arxiv.org/abs/2206.01338>

Song, Y., Lukasiewicz, T., Xu, Z., & Bogacz, R. (2020). Can the brain do backpropagation?---exact implementation of backpropagation in predictive coding networks. Advances in neural information processing systems, 33, 22566-22579. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7610561/>

Shen, G., Zhao, D., & Zeng, Y. (2022). Backpropagation with biologically plausible spatiotemporal adjustment for training deep spiking neural networks. Patterns, 3(6). <https://www.sciencedirect.com/science/article/pii/S2666389922001192>

Jeong, H., Taylor, A., Floeder, J. R., Lohmann, M., Mihalas, S., Wu, B., ... & Namboodiri, V. M. K. (2022). Mesolimbic dopamine release conveys causal associations. Science, 378(6626), eabq6740. <https://www.science.org/doi/10.1126/science.abq6740>

# Introduction

Highlights from our Gregynog poster 2024

* Backpropagation (BP) has been criticized for its lack of biological plausibility (Lillicrap et al., 2020). Contrastive Hebbian Learning (CHL) is an alternative that calculates error signals locally at each unit (Detorakis et al., 2019).
* CHL assumes a symmetry in the feedback weights matrix, a novel solution with random feedback improved CHL (rCHL; (Shervani-Tabar & Rosenbaum, 2023).
* Neither BP nor rCHL account for category learning (**Table 1** and **Table 2**). To solve this, we incorporated two rules: Mackintosh-like () and Pearce-Hall-like () to control different layers in BP and rCHL architecture (**Figure 1B**).
* and control the learning rate () change for each trial:  (Eq. 1; Pearce, et al., 1984; **Figure Simulations**).   
   is the average of absolutes prediction errors (Output-Target).
* We proposed rCHL with dynamic (rCHL-D). This algorithm solves category learning by slowly changing connections between Hidden-Output (H-O) relative to Input-Hidden (I-H; **Figure Simulations Weights**).
* Categories are encoded in I-H layer and at the reversal, networks only change the H-O connections strictly related with the ‘response’ (**Figure Simulations**).
* Implementing (Mackintosh-like(Mackintosh, 1975)) to I-H and (Pearce-Hall-like(Pearce & Hall, 1980)) to H-O both BP-D and rCHL-D solve category learning, but the later algorithm is more parsimonious with neurobiology.

## Modeling

A graph of a function

Description automatically generated with medium confidence

A chart of a diagram

Description automatically generated

# Experiment 1a – Easy Version

## Methods

### Participants

Thirty-two participants were recruited through Prolific ([www.prolific.com](http://www.prolific.com)). The age ranged between 20 and 63 with a mean of 36.29 and standard deviation of 10.86. Twenty participants were male, 7 female, one non-binary, and 4 did not respond.

We exclude participants that did not show adequate learning in the first 96 trials of the learning phase (before reversal, see below). We define adequate learning when the probability of correct categorical classification rejects the null under a binomial test where the probability of correct is higher than correct by chance, *i.e.*, 0.5. Thus, participants must have at least 57 correct trials over 96 trials in phase 1.

### Task and Procedures

The experiment consisted in two tasks: total- and partial-reversal. Every participant performed the two tasks. The order of the tasks was counterbalanced. Each task consisted in two phases: learning and reversal. The tasks had 6 blocks by phase. Each block consisted in 8 trials, 2 for each one of the trial types or fractals. Within each block the trials were fully randomized. Then, 6 (blocks) \* 8 (trials) \* 2 (phase) = 96 trials for each task.

Participants were asked to classify four fractals into two categories “North Hemisphere” (N) and “South Hemisphere” (S; see exact instructions in the Appendix). In every trial, participants saw one fractal exemplar and selected which category it belonged by pressing the left or the right keys (**Figure 1A**). Left was linked with the N (*i.e.*, N←; see **Table 1**) and right with S (*i.e.*, →S). Between trials there was an Inter-Trial Interval (ITI) with a white fixation cross and grey background. The ITI was a random variable distributed uniformly between .5 and 1.5 seconds rounded to the first decimal place.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 1.** | | | | | | | | | |
| Categorical Reversal Learning Experimental Design (easy) | | | | | | | | | |
| **Category Total Reversal** | | | | | **Category Partial Reversal** | | | | |
| Phase 1 | | Phase 2 | |  | Phase 1 | | Phase 2 | |  |
| Learning | | Reversal | |  | Learning | | Reversal | |  |
| Cues | Cat. | Cues | Cat. |  | Cues | Cat. | Cues | Cat. | Condition |
| A | N← | A | →S |  | A | N← | A | →S | reversal |
| B | N← | B | →S |  | B | N← | B | N← | non-reversal |
| C | →S | C | N← |  | C | →S | C | N← | reversal |
| D | →S | D | N← |  | D | →S | D | →S | non-reversal |
| *Note*: A, B, C, and D are cues in form of fractals; S← and →N are two distinct categories (Cat.). | | | | | | | | | |

### Software and Open Science

The task was programmed in PsychoPy® (Peirce et al., 2019) and the experiment was run online via Pavlovia in Amazon Mechanical Turk®. All the statistical analysis were conducted in R (R Core Team, 2021). The PsychoPy code and R scripts are available in GitHub, along with the analysis scripts and previous drafts of the paper: <https://github.com/santiagocdo/categoricalReversalLerning>

All the fractals used as stimuli in this experiments were kindly provided by: <https://www.enchgallery.com/index.htm>

### Analysis

To test the hypothesis, we used only the experimental data for the second phase, i.e., post-reversal. We modelled correct category classification (correct=1, incorrect=0) with a Logistic Mixed Model (LogMM):

, (*eq. 1*)

The previous model contains Blocks as random slope and participants as random intercept. Then to confirm whether the effect is due to the reversal, we ran another LogMM with only data from the first block in the second phase.

, (*eq. 2*)

No random slope was used, and participants were random intercept. Finally, to model Reaction Time (RT), we assumed that RT was normally distributed, and we used a Linear Mixed Model (LinMM), equivalent to *eq. 1*:

, (*eq. 3*)

Same random structure as *eq. 1*. The significance alpha for all models’ parameters was set to 0.05.

A screenshot of a computer

Description automatically generated

**Figure 1**. **A**. present one trial for both tasks used in Experiments 1a and 1b, and 2. **B**. Architecture used with all models: BP; CHL; rCHL; and rCHL-D𝛼 (Castiello et al., 2021; Delamater, 2012). 𝜌 control 𝛼 changes from Input-Hidden, and 𝜇 controls 𝛼 changes between Hidden-Outputs. Hidden layer always consisted in four multimodal units (*i.e.*, fully connected).

## Results

### Faster Learning After Reversal in Total Condition

The mean of the probabilities of correct category classification for all blocks in both tasks are presented in **Figure 2A**. On one hand, the total task is represented by one line (red) given all examples from both categories experienced the same reversal condition in phase 2. On the other hand, the partial task has two conditions: reversed (blue) and nonreversed (green). The critical comparison is to compare in phase 2 the reversed exemplars between the total and the partial task (**Figure 2C**). The LogMM reveal a positive main effect for total vs partial [], suggesting that in phase 2, in average participants are more correct in the total task. Also, there is a positive main effect in blocks [], which suggest that that participants are learning across trials. However, the interaction between task and blocks was not significant []. Just to corroborate that partial reversal impaired reacquisition of the categories, we tested the first block between total vs partial, and we found a significant and large effect [].

### No Differences in Reaction Time After Reversal

Total-Condition []

Block []

Interaction []

A group of graphs with different colored lines

Description automatically generated

**Figure 2**. Left panels Experiment 1, right panels Experiment 2. **A** and **D** probability of correct as a function of blocks. **B** and **E** mean of median reaction times as a function of blocks. **C** and **F** same as **A** and **D** but only plotted for the reversal part of the task, after the vertical line between bocks 6th (easy task) and 8th (hard task). We also removed the partial-nonreversed (green colour) conditions. On this data we tested the hypothesis (see stats). Error bars are standard errors of the means.

# Experiment 1b – Hard Version

## Methods

### Participants

Thirty-two participants were recruited through Prolific ([www.prolific.com](http://www.prolific.com)). The age ranged between 22 and 65 with a mean of 40.16 and standard deviation of 13.80. Eleven participants were male, 15 female, one non-binary, and 5 did not replied.

We exclude participants using the same criteria as Experiment 1a. Thus, participants’ inclusion is based on a binomial test in which the probability of correct should be higher than random correct 0.5. This implies at least 142 correct trials over 256 trials in phase 1.

### Task and Procedures

The task structure is the same as Experiment 1a. The only difference in this experiment is the difficulty which was manipulated by adding more stimuli (fractals; **Table 2**).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2.** | | | | | | | | | |
| Categorical Reversal Learning Experimental Design (hard) | | | | | | | | | |
| **Category Total Reversal** | | | | | **Category Partial Reversal** | | | | |
| Phase 1 | | Phase 2 | |  | Phase 1 | | Phase 2 | |  |
| Learning | | Reversal | |  | Learning | | Reversal | |  |
| Cues | Cat. | Cues | Cat. |  | Cues | Cat. | Cues | Cat. | Condition |
| A | S← | A | →N |  | A | S← | A | →N | reversal |
| B | S← | B | →N |  | B | S← | B | →N | reversal |
| C | S← | C | →N |  | C | S← | C | S← | non-reversal |
| D | S← | D | →N |  | D | S← | D | S← | non-reversal |
| E | →N | E | S← |  | E | →N | E | S← | reversal |
| F | →N | F | S← |  | F | →N | F | S← | reversal |
| G | →N | G | S← |  | G | →N | G | →N | non-reversal |
| H | →N | H | S← |  | H | →N | H | →N | non-reversal |
| *Note*: A, B, C, D, E, F, G, and H are cues in form of fractals; S← and →N are two distinct categories (Cat.). | | | | | | | | | |

### Analysis

Same analysis plan as Experiment 1a.

## Results

### Faster correctness after reversal in total condition

All model

Total-Condition []

Block []

Interaction []

First block

Total-Condition []

### No differences in reaction time after reversal

All model

Total-Condition []

Block []

Interaction []

# Experiment 2 – Stimuli Control

## Methods

### Participants

Sixteen participants were recruited through Prolific ([www.prolific.com](http://www.prolific.com)). The age ranged between 28 and 61 with a mean of 42.69 and standard deviation of 9.91. Six participants were male and 7 were female.

### Task

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 3.** | | | | | | | | | | |
| Categorical Reversal Learning Experimental Design (easy) | | | | | | | | | | |
|  |  | Right | | | | | | | | |
|  |  | A | B | C | D | E | F | G | H |
| Left | A |  | x | x | x | x | x | x | x |
| B | x |  | x | x | x | x | x | x |
| C | x | x |  | x | x | x | x | x |
| D | x | x | x |  | x | x | x | x |
| E | x | x | x | x |  | x | x | x |
| F | x | x | x | x | x |  | x | x |
| G | x | x | x | x | x | x |  | x |
| H | x | x | x | x | x | x |  |  |
| *Note:* 56 fractals similarity trials and ratings. 8 x 8 = 64, but 64 - 8 = 56, given the 8 stimuli diagonal, where it does not make sense to compare the same stimulus against itself. However, the comparison Left H and Right G was not tested due to a programming mistake. | | | | | | | | | | |

### Procedures

### Analysis

## Results

A group of graphs with numbers

Description automatically generated with medium confidence

# General Discussion

# References

Castiello, S., Zhang, W., & Delamater, A. R. (2021). The retrosplenial cortex as a possible “sensory integration” area: A neural network modeling approach of the differential outcomes effect in negative patterning. *Neurobiology of Learning and Memory*, *185*, 107527. https://doi.org/10.1016/J.NLM.2021.107527

Delamater, A. R. (2012). On the nature of CS and US representations in Pavlovian learning. *Learning & Behavior*, *40*(1), 1–23. https://doi.org/10.3758/s13420-011-0036-4

Detorakis, G., Bartley, T., & Neftci, E. (2019). Contrastive Hebbian learning with random feedback weights. *Neural Networks*, *114*, 1–14. https://doi.org/10.1016/J.NEUNET.2019.01.008

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience 2020 21:6*, *21*(6), 335–346. https://doi.org/10.1038/s41583-020-0277-3

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276–298. https://doi.org/10.1037/h0076778

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. https://doi.org/10.1037/0033-295X.87.6.532

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/S13428-018-01193-Y/FIGURES/3

R Core Team. (2021). *R A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. httpswww.R-project.org

Shervani-Tabar, N., & Rosenbaum, R. (2023). Meta-learning biologically plausible plasticity rules with random feedback pathways. *Nature Communications*, *14*(1), 1–12. https://doi.org/10.1038/s41467-023-37562-1

# Appendix

## Instructions Experiment 1a and 1b

In this experiment, you will be presented with a series of abstract images that represent the molecular structure of various natural objects. Your task will be to learn from which of two regions in the world these objects come (Northern Hemisphere or Southern Hemisphere).

More specifically, you will see 1 of 8 different abstract images at a time and be asked to indicate whether you think that image reflects an object taken from the Northern or Southern Hemisphere. Choose the Left Arrow Key for Northern or the Right Arrow Key for Southern Hemisphere. At first, you will need to guess, but you will be provided with feedback after your answer to help you learn which objects come from Northern or Southern Hemispheres.

Your response times are also important. Please make your response choices as quickly, but also as accurately, as you can. Your feedback will display the time (in sec) that it took for you to reach your decision, and also if your choice was correct (with a high pitch sound) or not (low pitch sound).

There will be a break halfway through.

Press the space bar when you are ready to begin.

## Instructions Experiment 2

[categorySimilarity]

In this experiment, you will be shown pairs of abstract images. Your task is to rate how similar or different the two images appear to you.

The experiment consists of two parts. In the first part, you will see and rate 56 pairs of images, and in the second part, another 56 pairs. Below each pair, there will be a slider ranging from "very different" to "very similar." Use your mouse to select a position on the slider that best represents how similar or different you find the two images are relative to one another.

There will be a break between both parts.

Press the space bar when you are ready to begin.

[fractal\_similarity]

In this experiment, you will be shown pairs of abstract images. Your task is to rate how similar or different each pair appears to you.

The experiment consists of two parts. In the first part, you will see 56 pairs of images, and in the second part, another 56 pairs, individual images may repeat between pairs. Below each pair, there will be a slider ranging from "very different" to "very similar." Use your mouse to select a position on the slider that best represents how similar or dissimilar you find the images.

There will be a break between both parts.

Press the space bar when you are ready to begin.

## Informed Consent

This research investigates the psychological processes used when people learn to identify objects in the world. You are being asked to participate in this research study because you are a normal healthy adult, and we wish to better understand basic learning processes in your population. The purpose of this research is to gain more knowledge about the cognitive processes involved in simple forms of associative learning.

If you agree to participate, we will ask you to perform in a simple computer task that will last approximately 12 minutes. In this task, you will see a series of abstract images presented individually on the screen and your task will be to learn to choose one of two response options for each image. Also, you will be asked to respond quickly and accurately on your computer keyboard when the image appears.

• Risks/Discomforts: There are no risks for participating in this study beyond those associated with normal computer use over a 15 min period.

• Benefits: This research is not designed to directly benefit you, but your help with this study will advance basic science on the cognitive processes involved in predictive learning in normal healthy individuals. Ultimately, this research could lead to a better understanding of some of the associative learning processes that are negatively impacted by various psychological conditions (such as aging, dementia, etc).

• Confidentiality: This study does not collect identifying information, and all data collected will remain anonymous. We will ask about your gender, age, and nationality, and we will record your performance in the task itself. However, this information will not be linked directly to any individual participant. The data we obtain will be stored indefinitely and may be shared publicly via online repositories for findings that are ultimately published.

Your participation in this research is completely voluntary, and you will be able to stop at any time without penalty. If you have any questions, you can contact: Andrew R. Delamater (andrewd@brooklyn.cuny.edu). If you have any questions about your rights as a research participant or if you would like to talk to someone other than the researcher, you can contact CUNY Research Compliance Administrator at 646-664-8918 or HRPP@cuny.edu.

If you wish to participate in the study, please press the spacebar for additional instructions.