Project Report: Team Reward

Hariharan Kalimuthu
2020115015

***Abstract*—This report is a collection of a few hypothesis we made up and tested on the dataset collected that was collected from 151 participants who voluntarily participated in a behavioral research[1] conducted by Kate Nussenbaum and Catherine A. Hartley to assess how specificity of details learnt changes with reward statistics and age. The details about the experiments, dataset, and the analysis conducted by us along with the results and conclusions have been summarised in the following ppt which can be accessed by clicking here.**

I. Experiment

The experiment involved participants of all ages ranging from children to adults, and they were given two tasks: a reinforcement learning task, followed by a memory task that was conducted about a week later.

In the reinforcement learning tasks, participants were shown stimuli associated with a particular reward associated with(of positive or negative magnitude). They had to choose whether to avoid the stimuli or approach the same. If they approach the stimuli, the reward associated with the stimuli got added to the subjects' scores. The goal was to maximize their score. This task was done in two conditions(termed block conditions), where the rewards assigned to the stimuli either was decided on a categorical basis(i.e. the reward for the stimuli was decided based on a broad category the stimulus fit into) or individual basis based on the stimuli(henceforth referred to as block condition).

A week later, participants had to undertake a memory task where they were shown stimuli and they had to identify whether they had seen the shown stimuli earlier or it was a novel stimuli.

The block conditions were set in place to ascertain how individuals learn and remember the specificity of a specific stimuli and the effect of the reward mechanisms set in place that motivates them to do the same. This experiment had two parts where the second part of the experiment involved tweaking some aspects of the first experiment to further analyse some variable interactions.

II. Exploratory Dataset Analysis

The data collected from the two tasks was made available and in order to come up with hypothesis we could work on and test on this dataset, we first conducted an exploratory dataset analysis on the datasets, which can be accessed by clicking here.

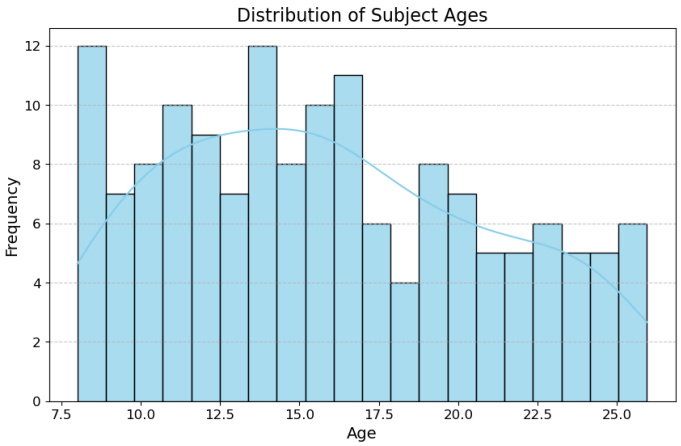

Fig. 1. Distribution of Subject Ages in the Experiment

The analysis gave us an idea of the ranges in which the variables lied and the to some extent the distribution of variables and how they might be correlated and/or interact with each other. The link to our GitHub repository along with the code to acccess the EDA as well as other parts of analysis can be accessed by clicking here.

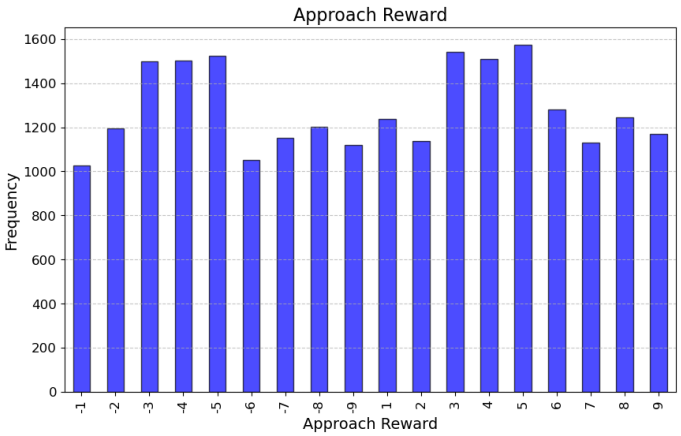

Fig. 2. Distribution of magnitudes of Reward

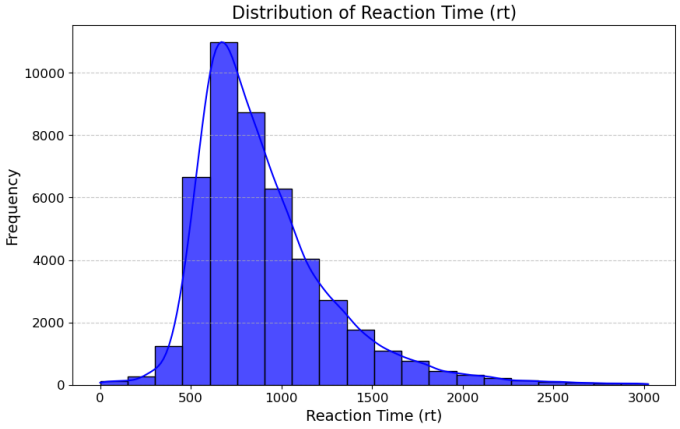

Fig. 3. Distribution of Reaction time by subjects

Hence, in order to proceed with our analysis, we came up with six hypothesis to test based on the dataset.

III. Hypothesis Testing

A. Specificity and Subsequent Memory Formation

H_0: There is no statistically significant difference in memory representations between individuals with different levels of specificity during learning.

H_1: There is a statistically significant difference in memory representations between individuals with different levels of specificity during learning.

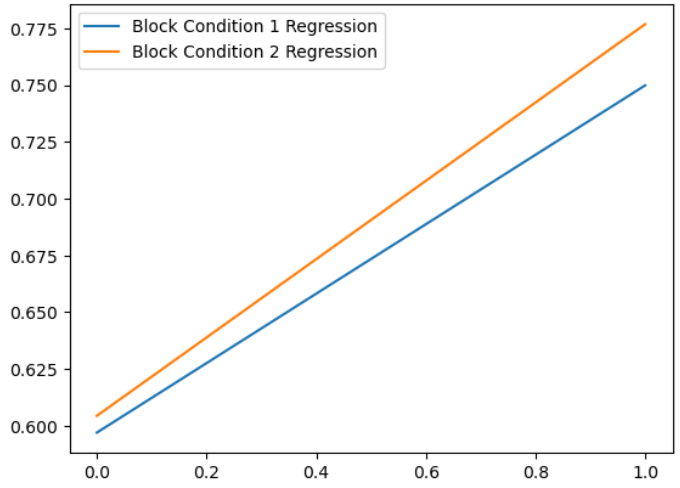

Fig. 4. Memory Scores and Weighted Confidence Levels

The testing for this hypothesis was done by my teammate Shreeya Singh. The analysis concluded that we could reject the null hypothesis and accept H1. Hence, the block condition significantly affects memory accuracy. More specific representations used during learning help the subjects remember the stimuli better and with higher confidence.

B. Age and Adaptability towards Specifity

Since the process of learning and remembering specific details is consuming in terms of cognitive resources, we tend to remember generalised information unless the situation necessitates specific details to be remembered. The hypothesis we worked on with respect to this aspect of memory is:

H_0: There is no significant relationship between age and the adaptation towards specificity in memory representations.

H_2: Adaptation towards specificity in memory representations increases/decreases with age.

This hypothesis was tested by my teammate Srujana Vanka and the conclusion to the same came out to be that Memory specificity adaptation increases with age, particularly in the Exemplar Predictive condition. While age has no significant impact on memory accuracy in Condition 1, it notably influences accuracy within Exemplar Predictive specificity. i.e. We accept the alternate hypothesis.

C. Age and Learning/Memory Specificity

In this part, we tried to study if the specifity of learning/memory increases with age. The hypothesis we generated and tested for the same was:

H_0: There is no significant relationship between age and the adaptation towards the specificity of learning computations.

H_3: Adaptation towards the specificity of learning computations increases/decreases with age.

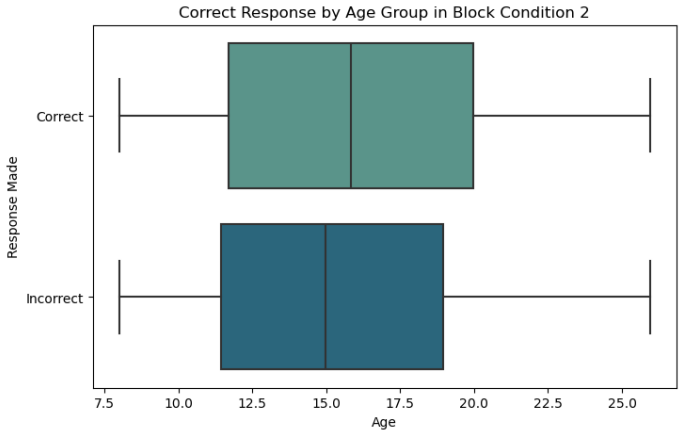

Fig. 5. Age vs Response Correctness

My teammate Srujana conducted hypothesis testing for this as well and the conclusion Adaptation towards learning specificity increases with age, impacting learning performance. Adults demonstrate higher correct response rates compared to younger age groups within each block condition. Hence, we rejected the null hypothesis and accepted the alternate hypothesis.

D. Examination of relations between stimulus repetition, reaction times and age

We conducted this analyses in two parts: First Shreeya examined the effects between reaction times and stimulus repetition. The hypothesis generated for the same is as follows:

H_0:There is no signification relationship between stimulus repetition and reactions times.

H_4a:Repeated stimuli have faster/slower reaction times.

As expected, the analysis concluded that the reaction time goes down with the repition of a particular stimulus, this effect can clearly be seen in the graph below:

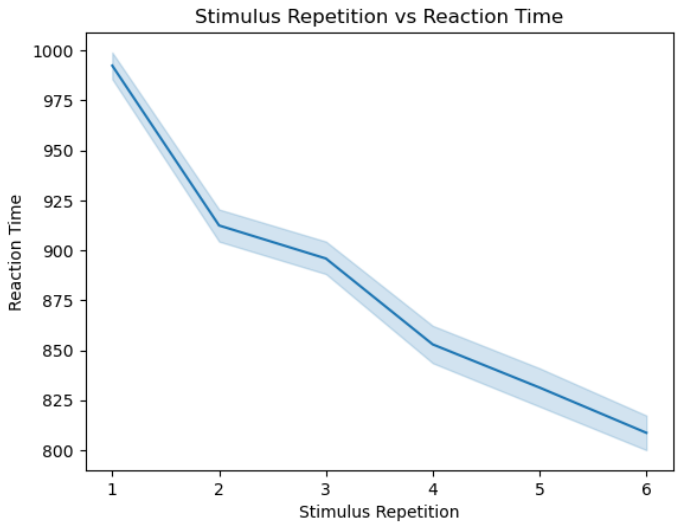

Fig. 6. Number of time a stimulus is shown vs avg. reaction time

The second part of the hypothesis aimed at studying if this observed effect increased/decreased with age.

H_0:There is no significant relationship between this effect and age.

H_4b:This effect increases/decreases with age.

We first analysed that the reaction times in general decrease with age as shown in this graph:

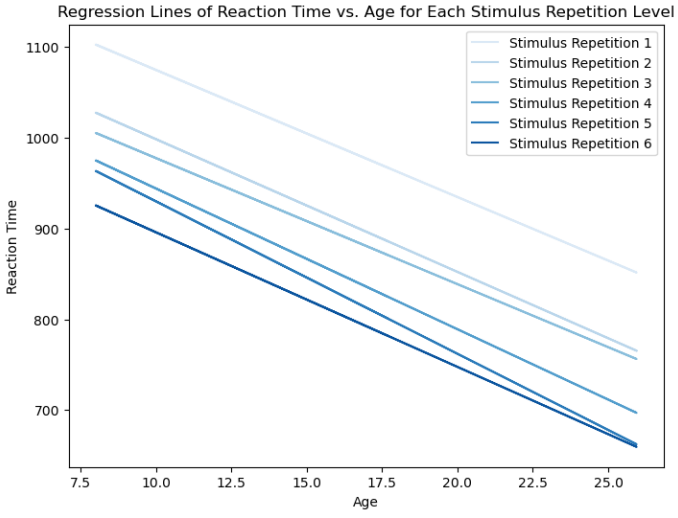

Fig. 7. Reaction Time vs Age, at various repetitions of stimuli

We conducted regression analysis to understand how these variables interact. We categorised the subjects into three groups 'Child', 'Adolescent' and 'Adult', and found out that even though the reaction times decrease with stimulus repetition and also age, the effect 'decrease in reaction time on stimulus repetition' decreases with age.

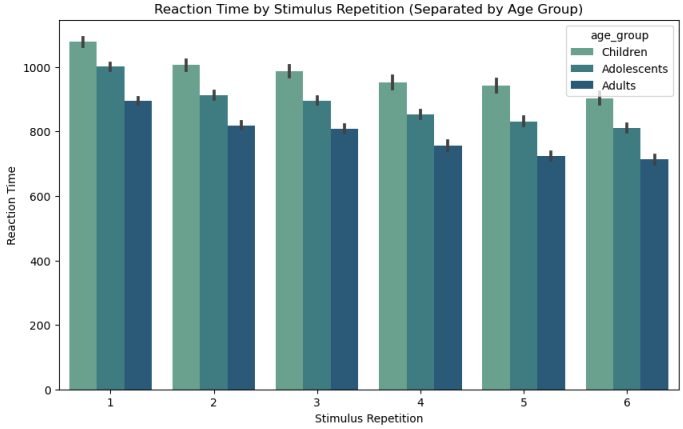

Fig. 8. Reaction Time vs Age Category and Stimulus Repition

E. Effect of Magnitude of Stimulus Reward

This part of the analysis which I have worked on, involved understanding whether the magnitude (whether positive/negative) of reward associated with the stimuli played a role in how well a subject learnt/remembered in the experiment. Hence, the hypotheses I set out with were:

On Learning and the Magnitude of Reward associated with the Stimulus.

H_0: There is no relation between the magnitude of reward associated with a particular stimulus and the learning about the stimulus.

H_5a: Higher the magnitude of the reward associated with a stimulus, better the subject learns to avoid/approach the stimulus.

On Memory and the Magnitude of Reward associated with the Stimulus.

H_0: There is no relation between the magnitude of reward associated with a particular stimulus and the memory associated with the stimulus.

H_5b: Higher the magnitude of the reward associated with a stimulus, better the subject remembers the stimulus.

For H_5a, we first checked if the variables involved(distribution of reward, points scored by the subjects) were homogeneous and normally distributed.

Shapiro-Wilk test(for Normality): For the distribution of the reward variable, the statistic came out to be 0.95 and p value was 0.79, and for the points obtained, the statistic came out to be 0.97 with p-value as 0.91. Hence, these do not significantly differ from a normal distribution.

Levene's test(for Homogeneity of Variances): The Levene's statistic came out to be 0.023 with p-value: 0.88. Hence, we can conclude that the variances of the two groups are more or less equal.

Effect of Magnitude of Reward on Learning: Since the distribution of variables was normal and homogeneous. We could use parametric tests for testing this hypothesis.

Pearson's Correlation: The value of the correlation statistic came out to be 0.24, with pvalue=0.23.

Hence, There is not enough evidence to conclude that there is a significant correlation between the variables in the population(p-value $>$0.05). We fail to reject the null hypothesis. Hence, there is no significant relation between learning and the magnitude of stimuli.

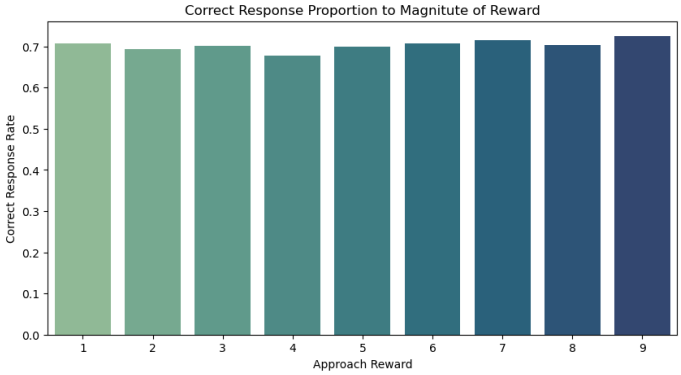

Fig. 9. Magnitude of Reward and Percentage Correctness of Responses for Learning Task

Now, similarly for the testing of hypothesis H_5b,

Shapiro-Wilk test(for Normality): For the distribution of the reward variable, the statistic came out to be 0.95 and p value was 0.72, and for the points obtained, the statistic came out to be 0.97 with p-value as 0.91. Hence, these do not significantly differ from a normal distribution.

Levene's test(for Homogeneity of Variances): The Levene's statistic came out to be 0.023 with p-value: 0.88. Hence, we can conclude that the variances of the two groups are more or less equal.

Effect of Magnitude of Reward on Memory: Since the distribution of variables is normal and they are homogeneous. We can use parametric tests for this hypothesis testing.

Pearson's Correlation: The value of the correlation statistic came out to be 0.25, with pvalue=0.35.

Hence, There is not enough evidence to conclude that there is a significant correlation between the variables in the population. We fail to reject the null hypothesis(as p-value $>$0.05).

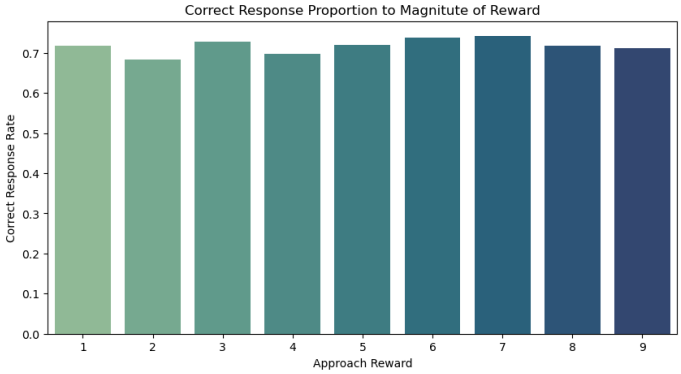

Fig. 10. Magnitude of Reward and Percentage Correctness of Responses for Memory Task

Hence, in this hypothesis testing, we understood that the magnitude of reward does not play a significant role in learning/remembering a stimuli well by the subject.

F. Learning, Specificity and Retention

In this part, we statistically try to prove a hypothesis that seems to be quite obvious at first glance. The research question we pose is simply: "Does Better Learning Lead to Better Retention? "

Hence, we can formally state the question posed above as the following hypothesis:

H_0: Points obtained by a subject in the learning task is unrelated to their performance in the memory task.

H_6: There is a positive relation between the magnitude of reward associated with a particular stimulus and the memory associated with the stimulus.

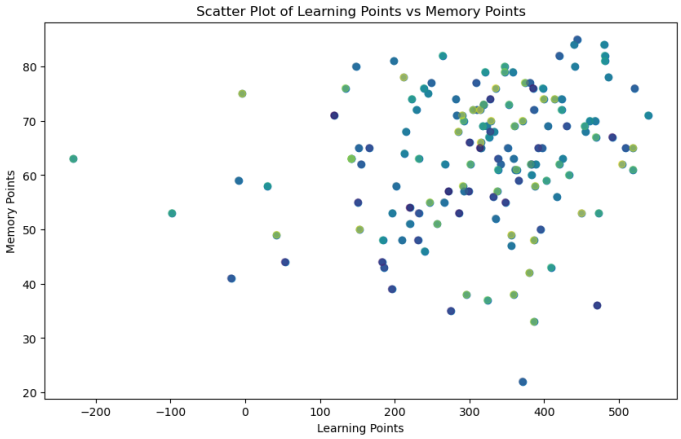

Fig. 11. Scatterplot of Points Scored by a Subject in Learning Task vs Memory Task

Shapiro-Wilk test(Test for Normality): For the distribution of the score for learning task, the statistic came out to be 0.94 and p value was 6.40 x 10e-6, and for the points obtained in memory task, the statistic came out to be 0.97 with p-value as 0.002. Hence, these variables significantly differ from normal distribution.

Levene's test(Test for Homogeneity of Variances): The Levene's statistic came out to be 141.29 with p-value: 5.99x10e-27. Hence, we can conclude that there are significant differences in variances between the groups or samples being compared.

Correlation between Learning and Memory:

Since the distribution of variables is not normal and they are not homogeneous. We need to make use parametric tests for this hypothesis testing.

Spearman's Rank Correlation: The value of the correlation coefficient came out to be 0.217, with p-value=0.007.

Hence, there is a statistically significant weak positive correlation between learning and memory. Hence, we can accept the alternate hypothesis and claim that better learning score leads to a better memory score.

Now, in order to understand how the block condition also impacts the memory score, we conducted a multiple regression analysis.

OLS Regression Results

Dep. Variable:	Memory_Scores_b1	R-squared:	0.072
Model:	OLS	Adj. R-squared:	0.059
Method:	Least Squares	F-statistic:	5.735
Date:	Sat, 04 May 2024	Prob (F-statistic):	0.00399
Time:	08:56:53	Log-Likelihood:	-587.14
No. Observations:	151	AIC:	1180.
Df Residuals:	148	BIC:	1189.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	58.0842	2.664	21.804	0.000	52.820	63.349
Learning_Scores_b1	-0.0096	0.016	-0.588	0.558	-0.042	0.023
Learning_Scores_b2	0.0412	0.013	3.285	0.001	0.016	0.066

Omnibus:	10.766	Durbin-Watson:	1.883
Prob(Omnibus):	0.005	Jarque-Bera (JB):	10.967
Skew:	-0.633	Prob(JB):	0.00416
Kurtosis:	3.376	Cond. No.	656.

Fig. 12. Multiple Regression Analysis

Based on the Regression Analysis, we see that:

Learning Scores Influence on Memory: Learning scores in block condition 2 significantly influenced memory scores, with a coefficient of 0.0412 (p = 0.001). In contrast, learning scores in block condition 1 did not show a significant effect on memory scores (coefficient = -0.0096, p = 0.558). Impact of Block Condition: The results suggest that the effect of learning performance on subsequent memory tasks may vary depending on the block condition in which the learning task was conducted. Specifically, Exemplar Stimuli leads to greater scores in the memory task compared to categorical stimuli. Overall Model Significance: The overall regression model was statistically significant (F-statistic = 5.735, p = 0.00399), indicating that the independent variables collectively explained a significant portion of the variance in memory scores. However, the model's adjusted R-squared value was low (0.059), suggesting that only a small proportion of the variance in memory scores was accounted for by the predictors included in the model.

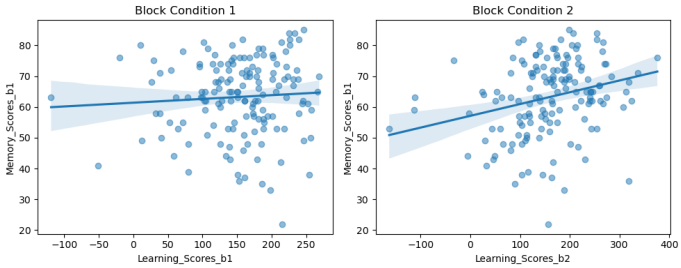

Fig. 13. Multiple Regression Analysis

IV. Conclusion

- Specificity-Memory Hypothesis: The block condition significantly affects memory accuracy. More specific representations used during learning help the subjects remember the stimuli better and with higher confidence.
- Age - Memory Specificity hypothesis: Memory specificity adaptation increases with age, particularly in the Exemplar Predictive condition. While age has no significant impact on memory accuracy in Condition 1, it notably influences accuracy within Exemplar Predictive specificity.
- Specificity in learning hypothesis: Adaptation towards learning specificity increases with age, impacting learning performance. Adults demonstrate higher correct response rates compared to younger age groups within each block condition.
- Stimulus Repetition with Reaction Times Stimulus Repetition-Reaction Times effect: Different stimulus repetitions indeed lead to varied reaction times, with some repetitions significantly affecting reaction speed compared to others. Stimulus Repetition-Reaction Times effect on age: Adults tend to exhibit faster reaction times as stimulus repetition increases, compared to children and adolescents.
- Effect of the magnitude of reward: A greater reward/punishment does not significantly affect learning/memory performance.
- Relation between learning, memory and specificity: Subjects who learn better, retain better. The effect is seen to be increased with increase in specificity of learning.

V. Division of Work

- **Hariharan Kalimuthu:** H_4b, H_5, H_6.
- **Shreeya Singh:** EDA, H_1, H_4a
- **Srujana Vanka:** EDA, H_2, H_3

References

[1] Kate Nussenbaum and Catherine A Hartley. *Reinforcement learning increasingly shapes memory specificity from childhood to adulthood*. Nov. 2023. DOI: 10.31234/osf.io/utsxn. URL: osf.io/preprints/psyarxiv/utsxn.