1 Results

1.1 Learner Characteristics

When asked the amount of time participants have used a chatbot- in any form or subject: 23 stated they had never used a chatbot, educational or not. Further, 18/42 having used a chatbot at least once for between 0-4 hours of use in total. Two individuals had spent much longer time with usage- these were a learning technologist associated with the CEPEH team, and a mature student. In short, this sample of participants did not regularly use chatbots for their course learning, but 1/3 had used one in some form previously.

Most learners use books or online books as resources. They may use multiple sources however they were asked to note the primary source. Only 6 stated their primary sources were *Online videos/interactive materials* which includes such tools as chatbots.

The first boxplot (1.1) shows learners perceptions of easy of use of mobile app and other educational mobile resources

```
## # A tibble: 3 x 83
##   sex    time  Code      Locat~1 Profe~2 Curre~3 Previ~4 UseCh~5 I fee~6 Usefu~7
##   <chr>  <chr> <chr>     <chr>   <chr>   <chr>   <chr>   <chr>   <chr>   <chr>
```

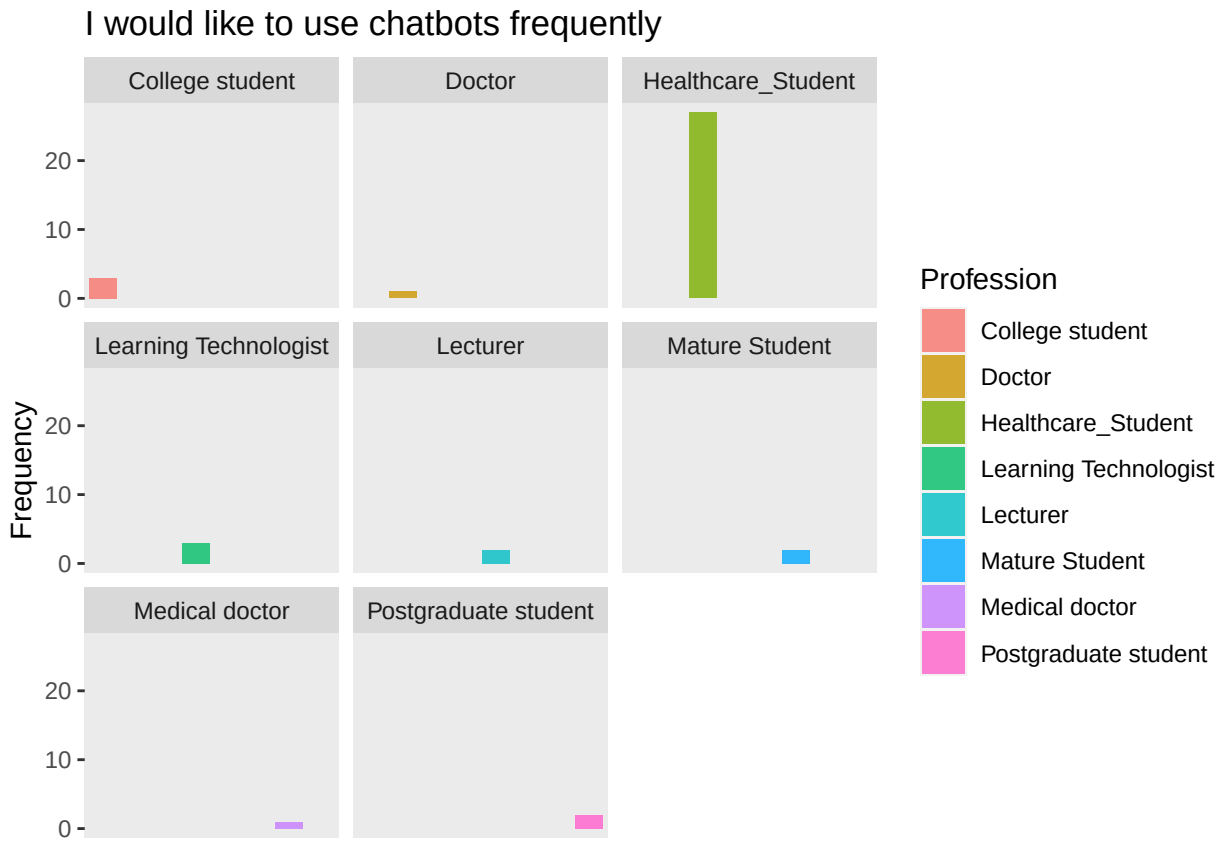


Figure 1.1: Usage Frequency-Pre

```
## 1 Female Pre   ER 31102~ Sweden  Studen~ One to~ Never   Agree   Agree   Agree
## 2 Female Pre   FAI2205   Greece  Studen~ Books   Never   Agree   Agree   Strong~
## 3 Female Pre   u8        Greece  Studen~ Books   1-4 ho~ Strong~ Agree   Agree
## # ... with 73 more variables: IncreaseAchievingPE2PRE <chr>,
## #   AccomplishThingsQuicklyPE3PRE <chr>, IncreasesProductivityPE4PRE <chr>,
## #   `HowToUse(EE1)(pre)` <chr>,
## #   `InteractionsClearUnderstandable(EE2)(pre)` <chr>,
## #   `EasyToUse(EE3)(pre)` <chr>, `EasySkilfulUsing(EE4)(pre)` <chr>,
## #   `PeopleImportantUse(SI1)(pre)` <chr>, `PeopleInfluenceUse(SI2)(pre)` <chr>,
## #   `PeopleOpinionsUse(SI3)(pre)` <chr>, ...
```

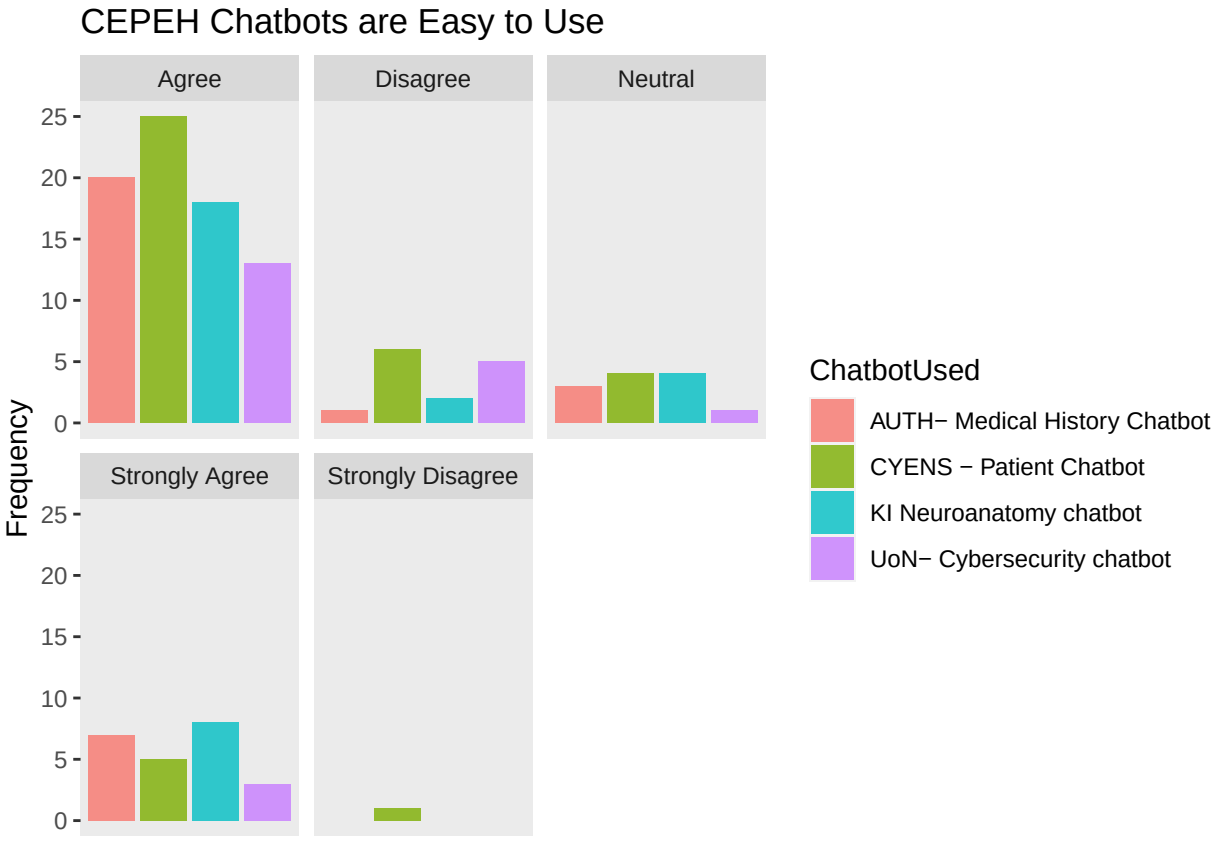
Previous_Chatbot_Usage	n
1-4 hours	15
10-19 hours	1
20+ hours	1

Previous_Chatbot_Usage	n
5-9 hours	2
Never	23

##Post-Intervention Results and Comparision

After using the CEPEH chatbots, majority of participants stated they would reuse the chatbots. However, there was 6 counts of *disagree* or *strongly disagree* for all 4 chatbots. Further, there were 17 counts of neutral in reuse, which was approximately 4 participants per chatbot (see (1.3).

For CYENS, even though the knowledge of the topic was not perceived to improve by some participants, this box plot shows how 34/42 stated they would reuse the chatbot developed by CYENS.



There was only 1 'Strongly Disagree' response. The agreement options counted for the majority of the data.

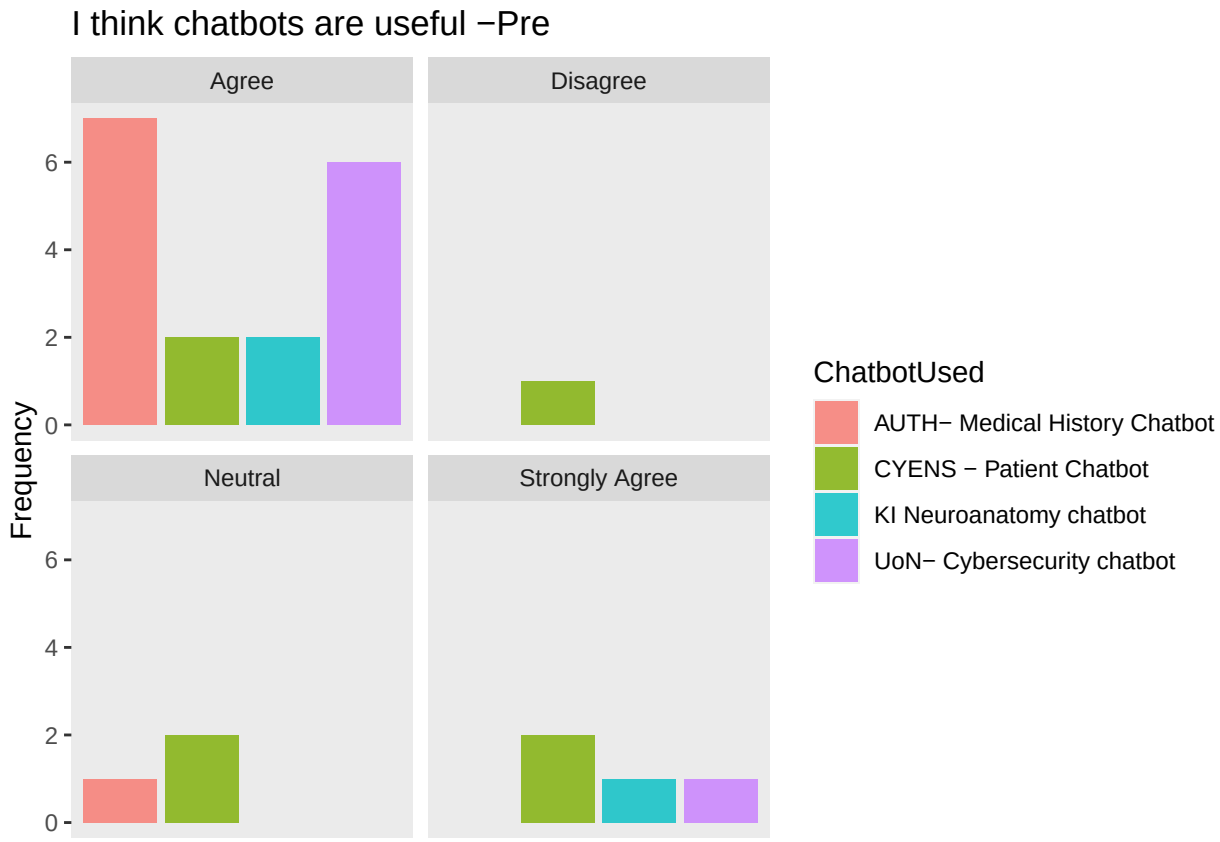


Figure 1.2: Useful Opinion-Pre

1.1.1 Other Findings

Other questions

I intend to continue using chatbots in the future (BI1)

The chatbot provided the information I needed with minimal commands

My knowledge of the topic improved after i had used the Chatbot

My confidence in understanding the topic improved after I had used the Chatbot

The chatbot provided me with the type of response i expected from asking a tutor/lecturer

The information provided was reliable

The chatbot has a high level of trustworthiness

The duration of conversations to find my answer was too long

The videos/images provided were useful to my questions

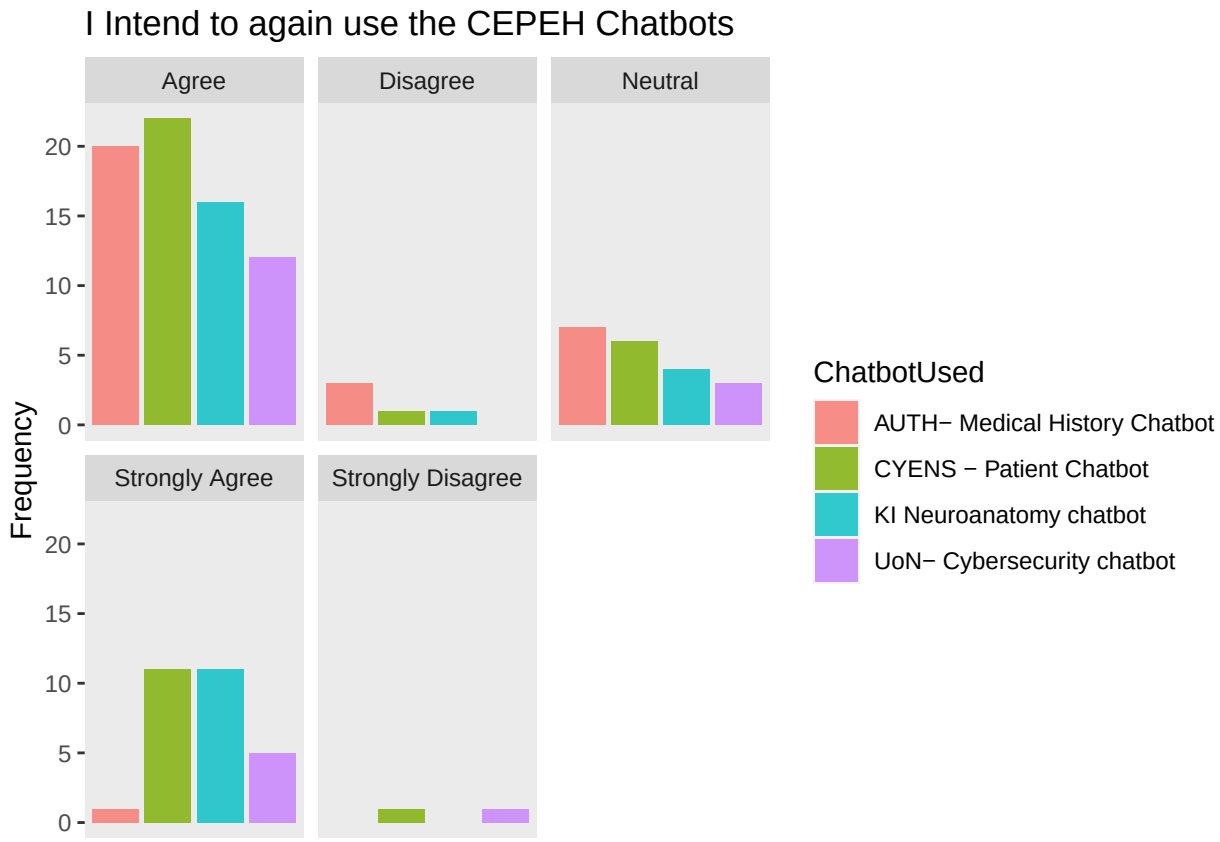


Figure 1.3: Intend to Reuse-Post

The chatbot exceeded my expectation of how it could help me

The chatbot exceeded my expectation of how it could engage with me

I think this learning method could help me to acquire knowledge

I would use this tool again as it has some value to me

I think i will actively use this learning method

I believe I had some choice about learning during chatbot use

I would trust the chatbot to provide me with information for my course

One piece of knowledge i learned from the chatbot was..

Repeated Measures t-test, aka paired t-test (before and after measurements)

This t-test compares confident using mobile chatbots before and after CEPEH chatbot usage.

1.1.2 System Usability Scale (SUS) Scores

Note= The amount of 'agreement' is defined as the addition of 'Agree' and 'Strongly agree' responses.

The SUS score should consist of 10 items. However, some SUS questions were improved upon by 1 or more CUQ questions, specifically to this Chatbot study. The SUS results would be overshadowed by the CUQ scores, expect 2 that did not have cross-over. The two questions were:

- I would like to use the CEPEH chatbot I tested, more frequently (SUS1)(post)
- I felt confident using the CEPEH chatbot (SUS2)(post)

This meant the score of the SUS was not created, however the CUQ score better represented the Learners' perceptions of the CEPEH chatbot in terms of feasibility of reuse and acceptability in healthcare curricula.

Keep_Using_Chatbots	Confident	Count
Agree	Agree	44
Agree	Disagree	5
Agree	Neutral	11
Agree	Strongly Agree	6
Disagree	Agree	6
Disagree	Disagree	5
Disagree	Neutral	4
Neutral	Agree	10
Neutral	Disagree	1
Neutral	Neutral	6
Not Applicable	Not Applicable	3
Strongly Agree	Agree	10
Strongly Agree	Not Applicable	1
Strongly Agree	Strongly Agree	12
Strongly Disagree	Agree	1
Strongly Disagree	Strongly Agree	1

1.1.3 Technology Acceptance Model

The TAM had 3 sections (Ease of Use, Perceived Usefulness, and Intention of Use). Ease of Use results showed significant increases in Users' usage with each Chatbot.

Perceived Usefulness: There were not significant findings for the Perceived usefulness. The justification for this may be due to being early versions of applications with limited functionality and functions which can be difficult for user to experience the intended further range of features and learning exercises.

Intention of Use: For users' intentions to use within their course, the result of the Mann-Whitney U test was not significant, U = , z = , p = . in their intentions before use (m=xx, mode=xx) compared to after (m=xx, mode=x), however there was improvement therefore the chatbots may have more benefit than expected by students.

1.2 Chatbot Usabilty Questionanire (CUQ)

1.2.1 CUQ Calcuation tool

The CUQ was developed by researchers at Ulster University,Link and as the calculation can be complex a dedicated calculation tool has been created. Please download the CEPEH CUQ calculation tool which has all of the data entered, so you can see the CEPEH CUQ scoring.

click here to download CUQ calc tool

click here to download CUQ score image *mobile download disabled

The score for all 3 chatbots grouped was 65.2/100, This scoring system was designed to be comparable to SUS and may be freely used alongside it, or in combination with other usability metrics. There has been evidence of correlation of 76% between the CUQ and SUS therefore we expect the SUS scored to be between 48.75 and 81%. We believe the CUQ has more validity towards measuring the concepts of interest on this study.

Read the CUQ development paper, see page 3 for correlation

Figure shows the CUQ scores as a box plot to highlight the range of Usability of the resources. Further exploration is required to understand which elements are causing this spread.

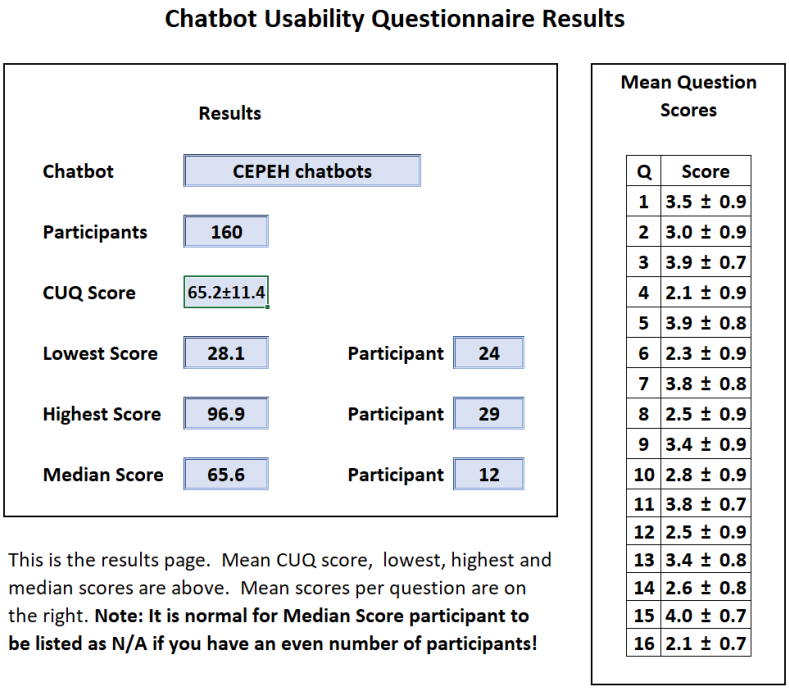
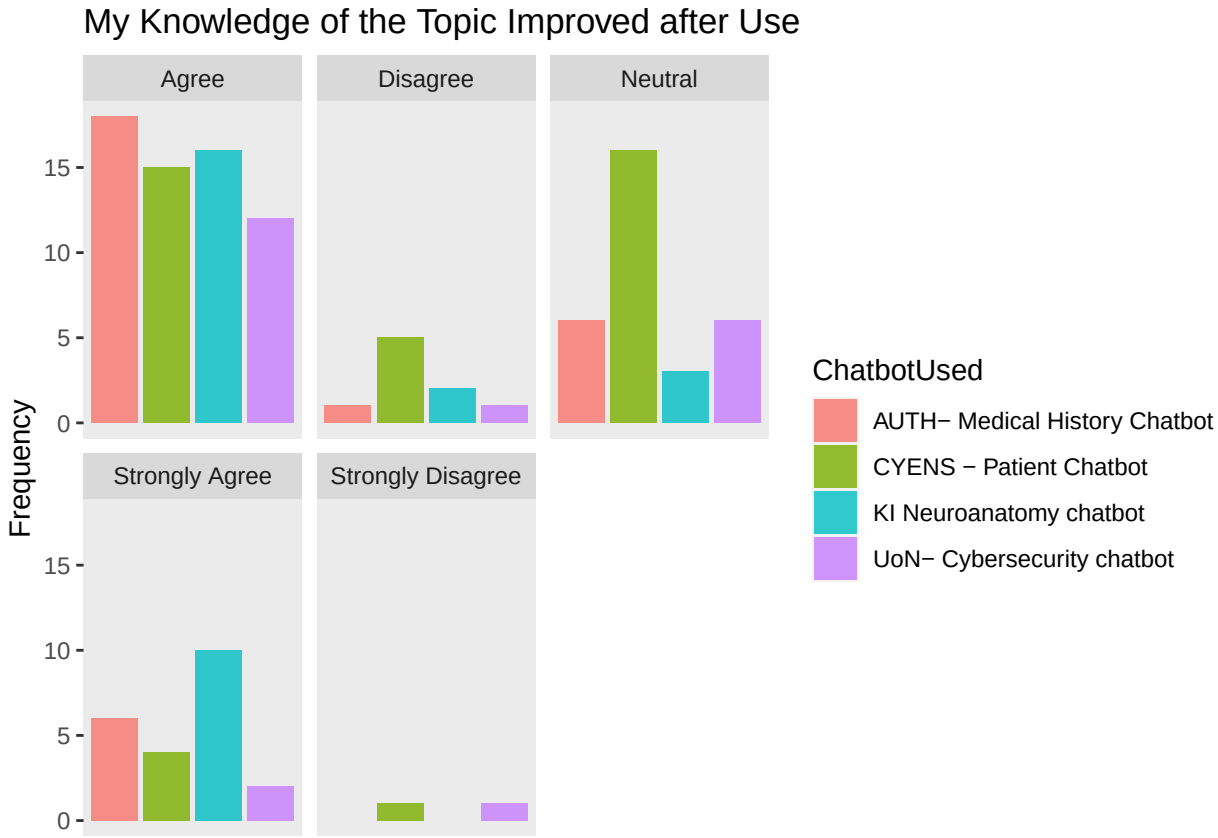


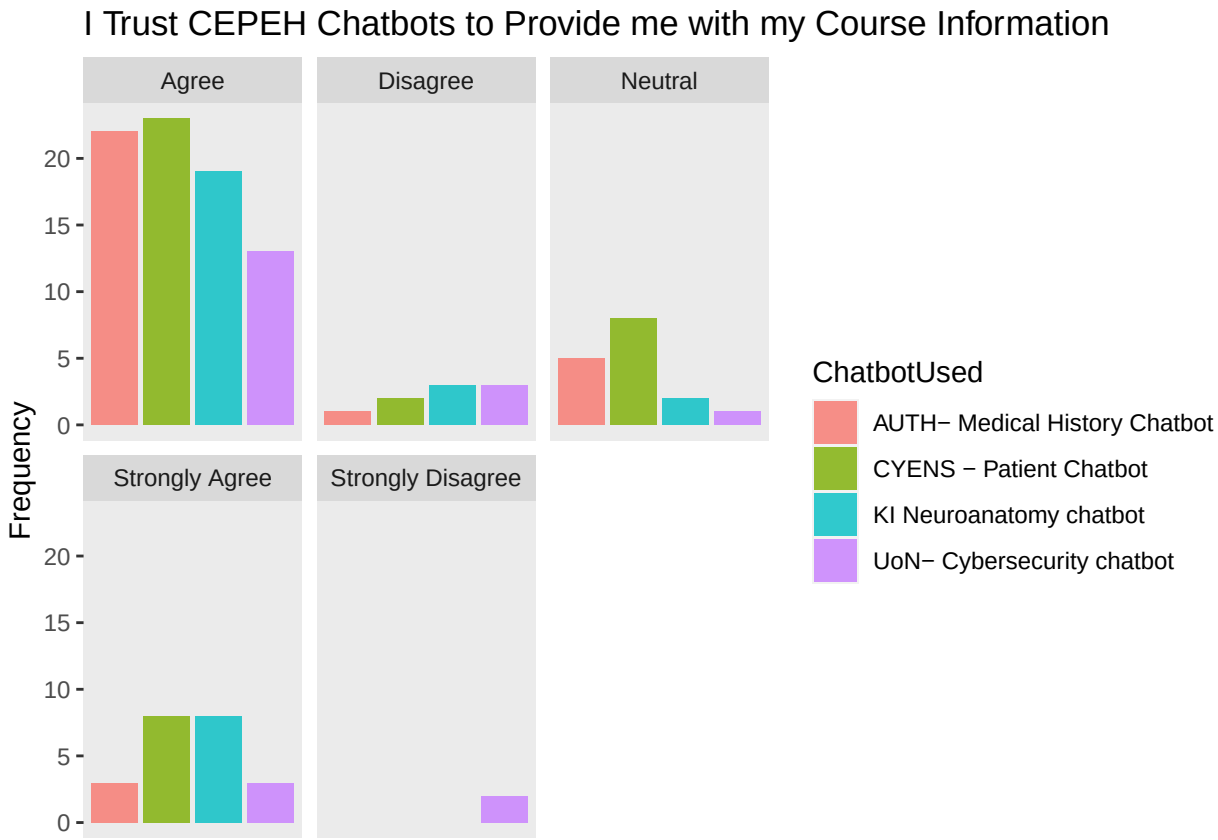
Figure 1.4: A marvel-lous meme

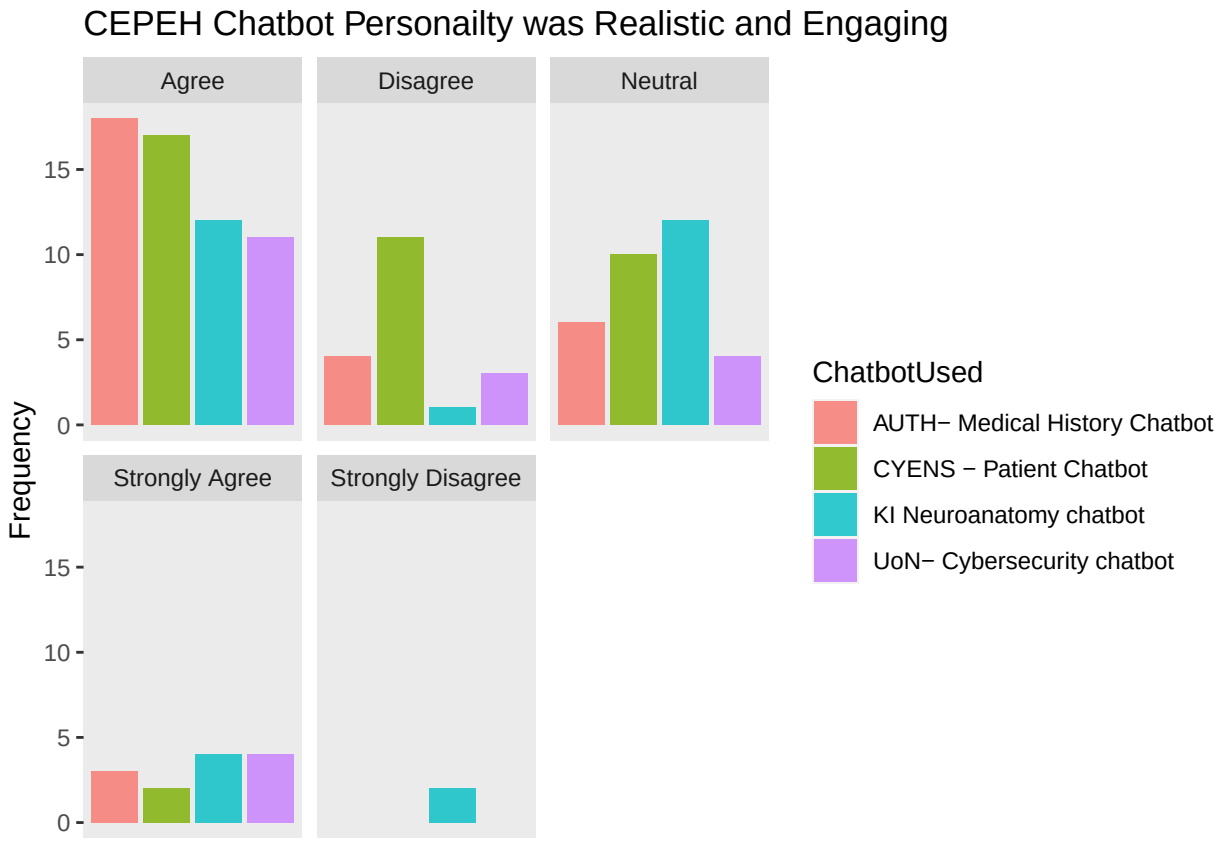
(#fig:cuq image)



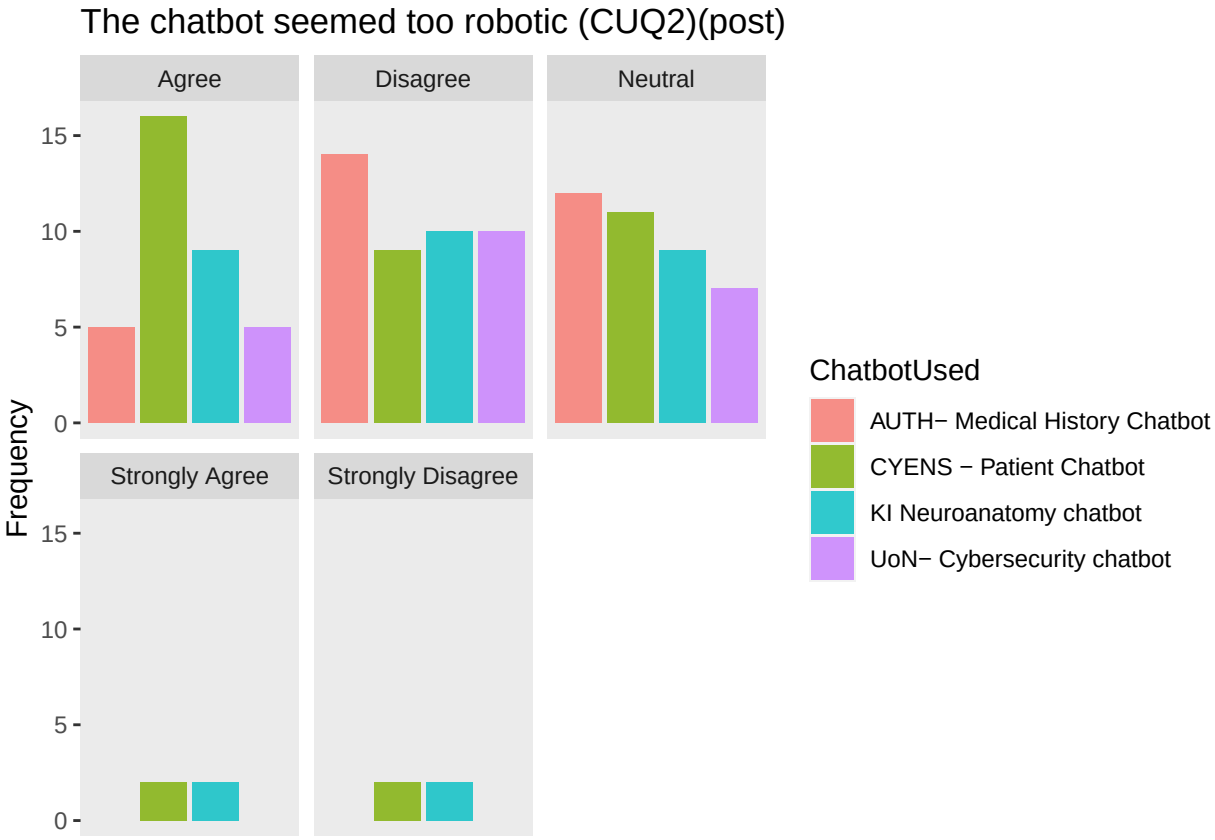
CYENS chatbot had around 10 more participants stating that they were neutral

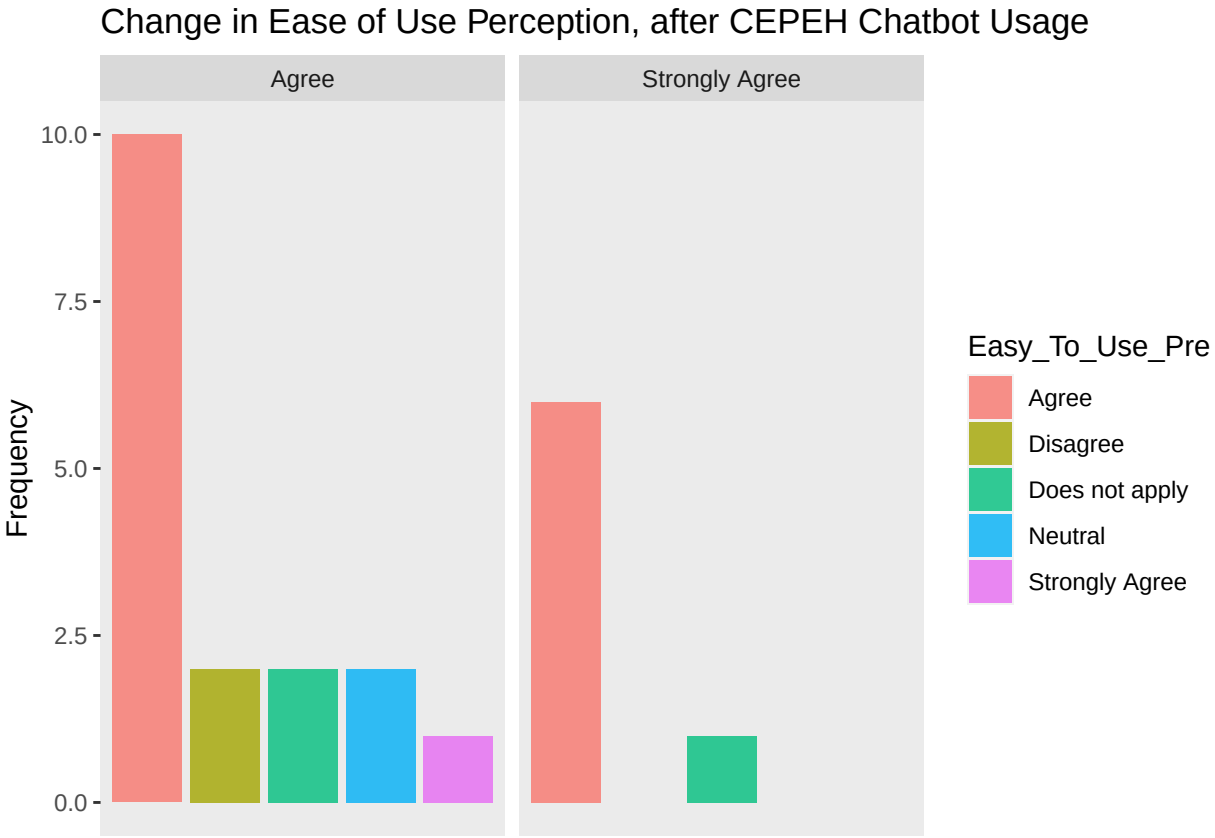
on gaining knowledge of the topic

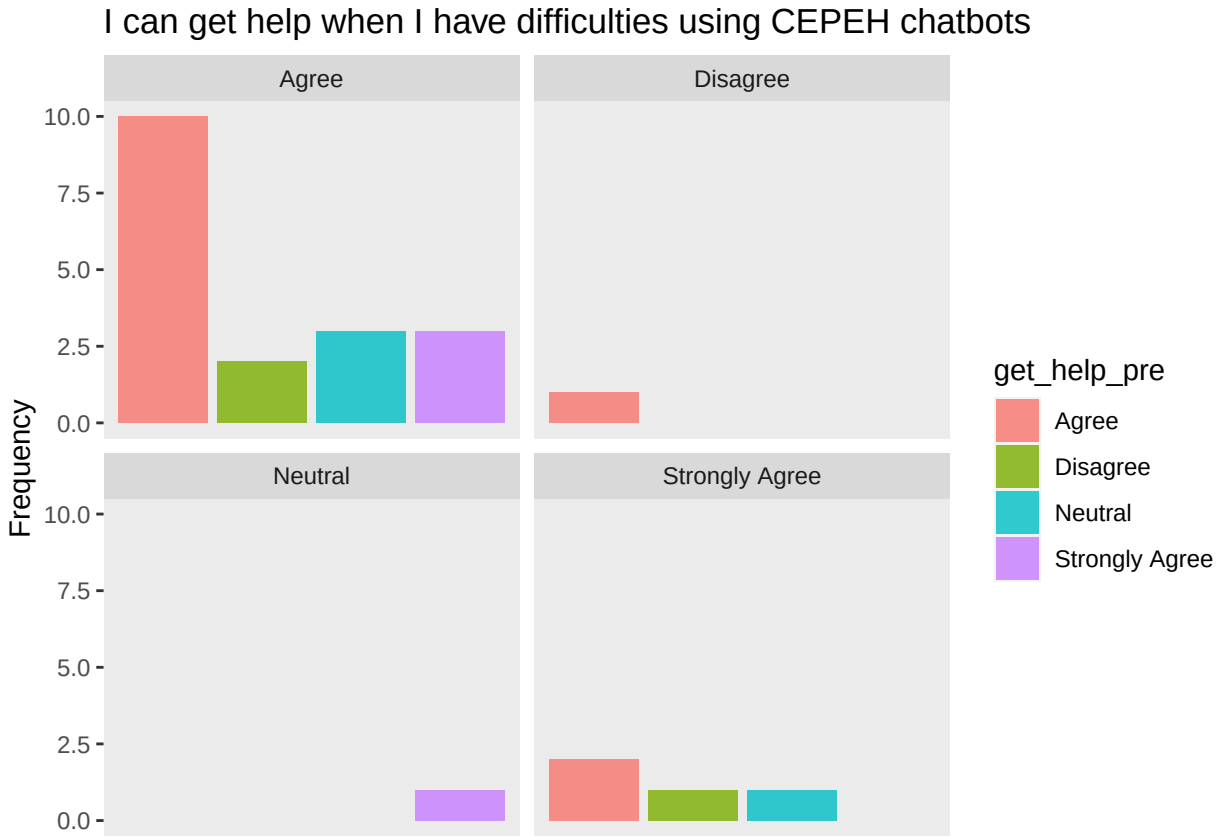




There was mixed results for the chatbot used being realistic and engaging. This question has two descriptive terms however based on the other results we understand that the chatbots' NLP logic, or ability to respond required improvement to be more 'smooth' in replying. The primary limitation was found in the 'robotic' interactions(See Figure 10). This was investigated further in the 'Text Mining' and 'Sentiment Analysis' sections.


The chatbot seemed too robotic (CUQ2)(post)
Agree
Disagree
Neutral
Strongly Agree
Strongly Disagree
Frequency
0
5
10
15
ChatbotUsed
AUTH− Medical History Chatbot
CYENS − Patient Chatbot
KI Neuroanatomy chatbot
UoN− Cybersecurity chatbot


Change in Ease of Use Perception, after CEPEH Chatbot Usage
Agree
Strongly Agree
Frequency
0.0
2.5
5.0
7.5
10.0
Easy_To_Use_Pre
Agree
Disagree
Does not apply
Neutral
Strongly Agree



Those who disagreed or were neutral in the pre usage measure, improved their understanding that help was available with the CEPEH chatbots. After usage, 40 participants agreed they could get help if they had difficulty using the resources.

1.3 Inferential Statistics