Comparing informativeness of an NLG chatbot vs graphical app in diet-information domain

Simone Balloccu and Ehud Reiter
Dept. of Computing Science, University of Aberdeen, UK

Horizon 2020
European Union Funding
for Research & Innovation

Communication by visualisation

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:
  - Visual deficit, information context, cognitive effort…

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:
  - Visual deficit, information context, cognitive effort…

- **Combining text and visualisations** proved to be effective in tackling this issue in:

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:
  - Visual deficit, information context, cognitive effort…

- **Combining text and visualisations** proved to be effective in tackling this issue in:
  - Healthcare (**van der Meulen et al., 2010**)

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:
  - Visual deficit, information context, cognitive effort…

- **Combining text and visualisations** proved to be effective in tackling this issue in:
  - Healthcare (**van der Meulen et al., 2010**)
  - Sensor data (**Molina et al., 2011**)

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:
  - Visual deficit, information context, cognitive effort…

- **Combining text and visualisations** proved to be effective in tackling this issue in:
  - Healthcare (**van der Meulen et al., 2010**)
  - Sensor data (**Molina et al., 2011**)
  - Weather data (**Gkatzia et al., 2017**)

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:
  - Visual deficit, information context, cognitive effort…

- **Combining text and visualisations** proved to be effective in tackling this issue in:
  - Healthcare (**van der Meulen et al., 2010**)
  - Sensor data (**Molina et al., 2011**)
  - Weather data (**Gkatzia et al., 2017**)
  - etc…

Communication by visualisation

- Insights are often delivered through visualization, but their **understanding can be hindered** by various factors:
  - Visual deficit, information context, cognitive effort…

- **Combining text and visualisations** proved to be effective in tackling this issue in:
  - Healthcare (**van der Meulen et al., 2010**)
  - Sensor data (**Molina et al., 2011**)
  - Weather data (**Gkatzia et al., 2017**)
  - etc…

- Past research inspected **static contexts** only (no interaction; fixed presentation).

Our contribution

Our contribution

- We implement an **NLG chatbot** that:

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**

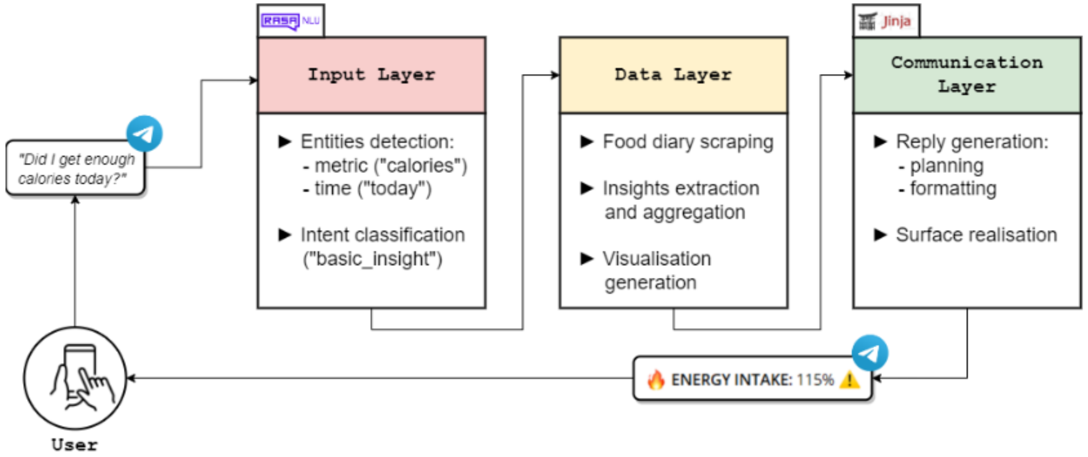


Figure 1: Chatbot architecture and interaction flow.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

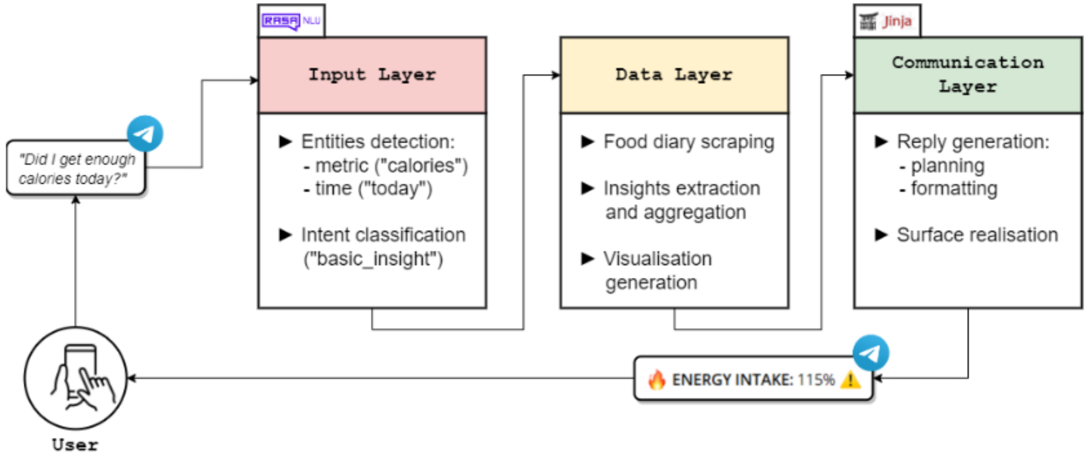


Figure 1: Chatbot architecture and interaction flow.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

- We adopt a hybrid architecture:

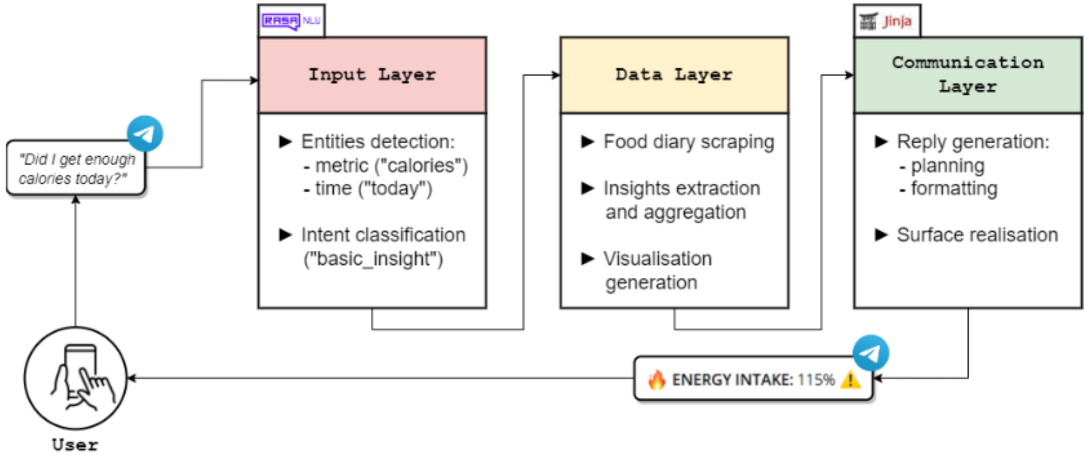


Figure 1: Chatbot architecture and interaction flow.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

- We adopt a hybrid architecture:
  - Machine-learning for **NLU**

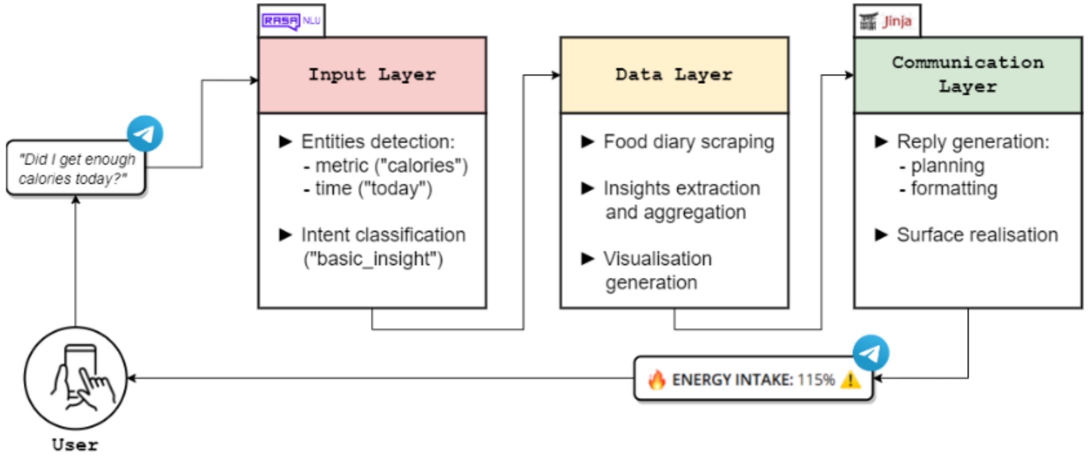


Figure 1: Chatbot architecture and interaction flow.

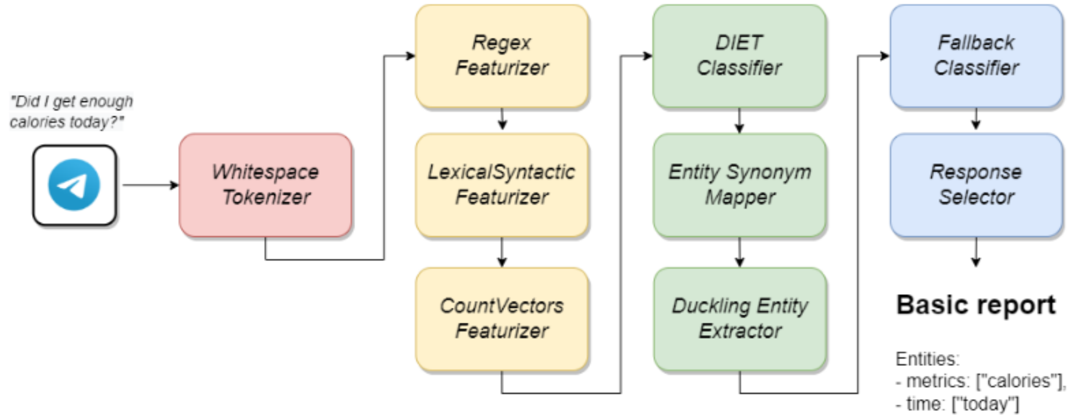


Figure 2: Overview of the NLU pipeline.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

- We adopt a hybrid architecture:
  - Machine-learning for **NLU**
  - Rule and templates for **NLG**

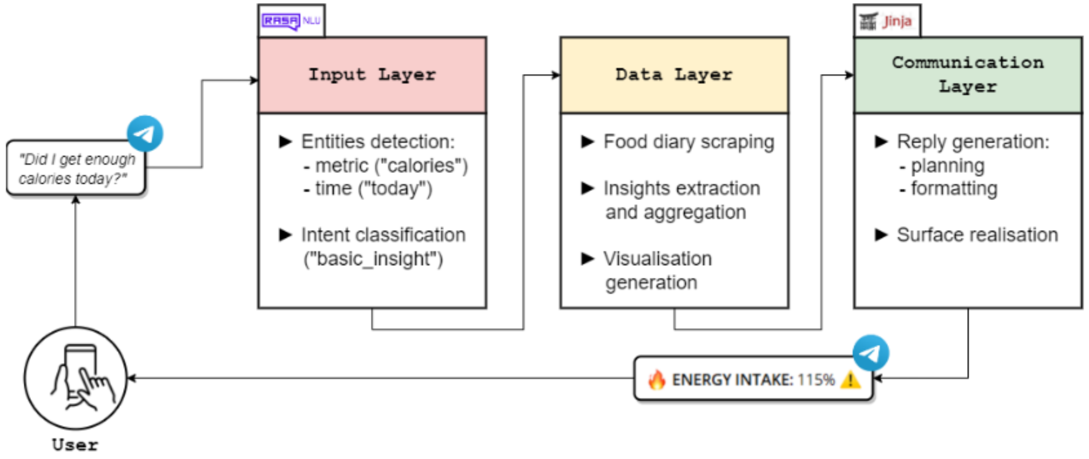


Figure 1: Chatbot architecture and interaction flow.

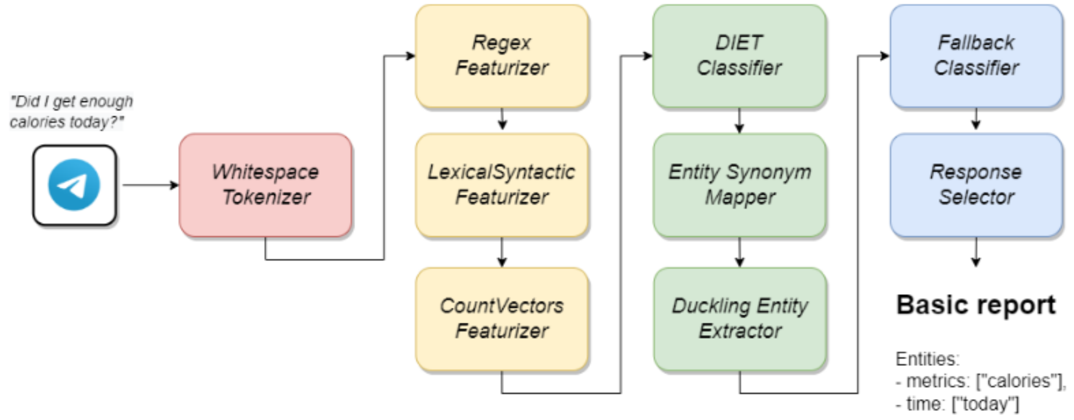


Figure 2: Overview of the NLU pipeline.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

- We adopt a hybrid architecture:
  - Machine-learning for **NLU**
  - Rule and templates for **NLG**

- We apply and evaluate it in the diet domain:

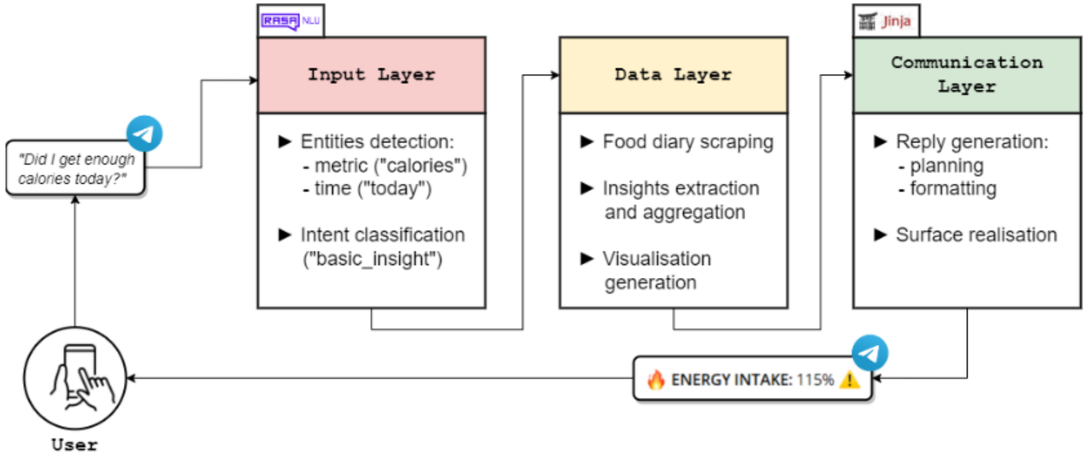


Figure 1: Chatbot architecture and interaction flow.

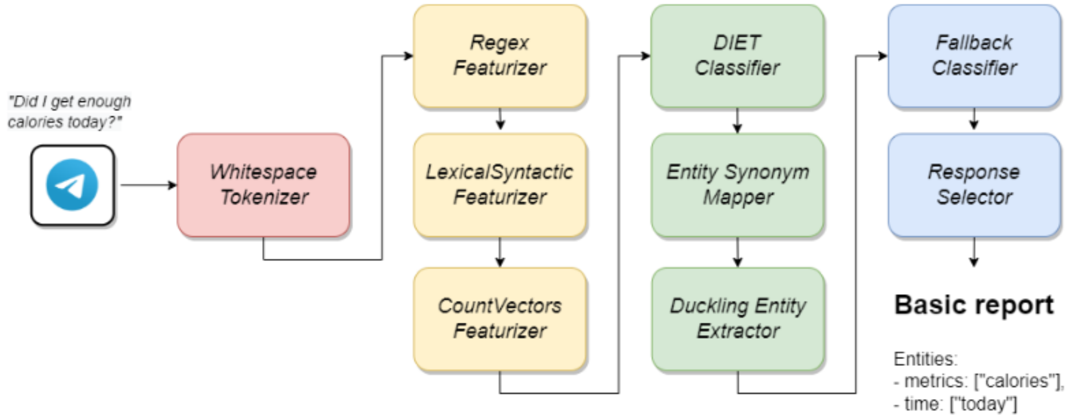


Figure 2: Overview of the NLU pipeline.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

- We adopt a hybrid architecture:
  - Machine-learning for **NLU**
  - Rule and templates for **NLG**

- We apply and evaluate it in the diet domain:
  - Information is critical (**Van Dorsten and Lindley, 2008; Savolainen, 2010; Michie et al., 2011**)

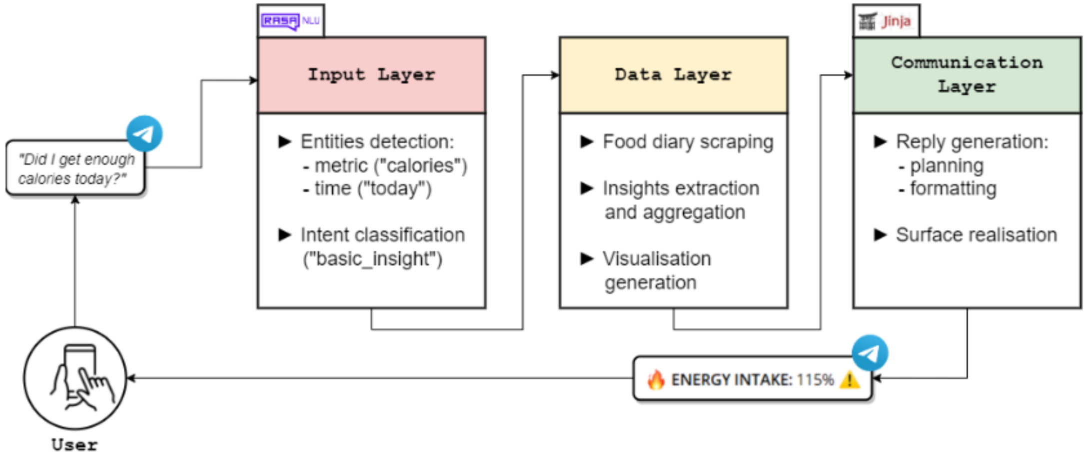


Figure 1: Chatbot architecture and interaction flow.

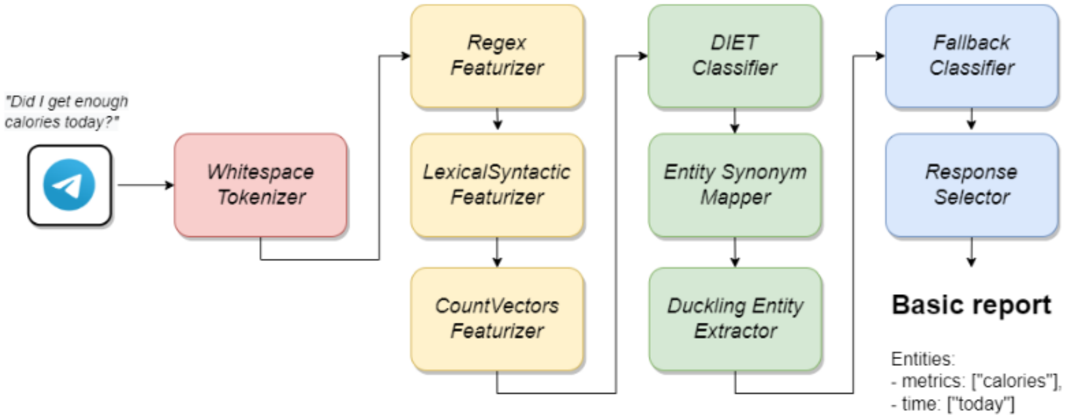


Figure 2: Overview of the NLU pipeline.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

- We adopt a hybrid architecture:
  - Machine-learning for **NLU**
  - Rule and templates for **NLG**

- We apply and evaluate it in the diet domain:
  - Information is critical (**Van Dorsten and Lindley, 2008; Savolainen, 2010; Michie et al., 2011**)
  - Efficient communication is a predictor of prolonged app use (**Lee and Cho, 2017**)

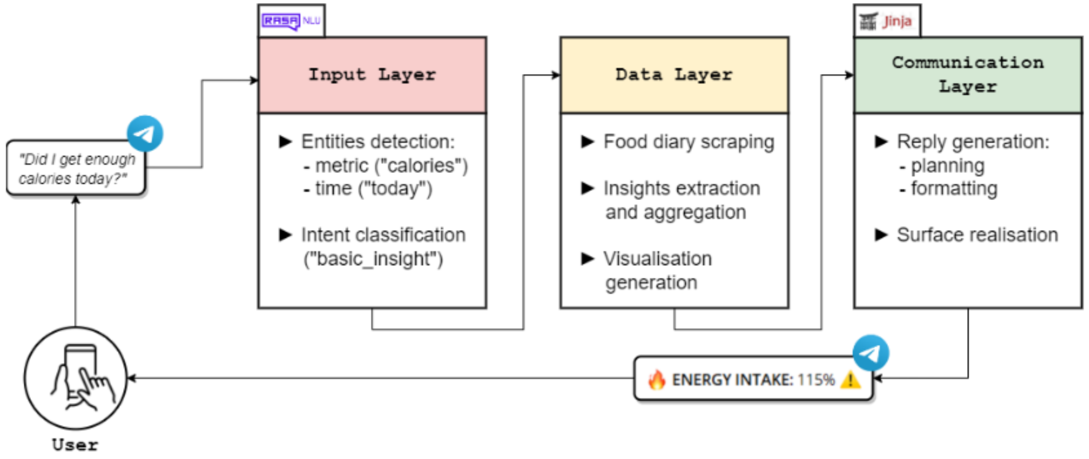


Figure 1: Chatbot architecture and interaction flow.

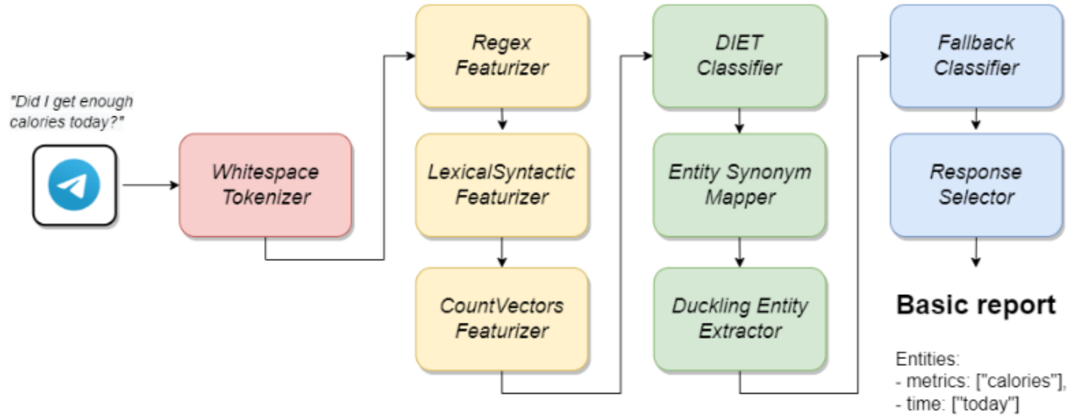


Figure 2: Overview of the NLU pipeline.

Our contribution

- We implement an **NLG chatbot** that:
  - Processes **natural language queries**
  - Provides insight explanation through **charts and text combination**

- We adopt a hybrid architecture:
  - Machine-learning for **NLU**
  - Rule and templates for **NLG**

- We apply and evaluate it in the diet domain:
  - Information is critical (**Van Dorsten and Lindley, 2008; Savolainen, 2010; Michie et al., 2011**)
  - Efficient communication is a predictor of prolonged app use (**Lee and Cho, 2017**)
  - Existing dieting apps adopt poor communication (**Balloccu et al., 2021; Balloccu and Reiter, 2022**)

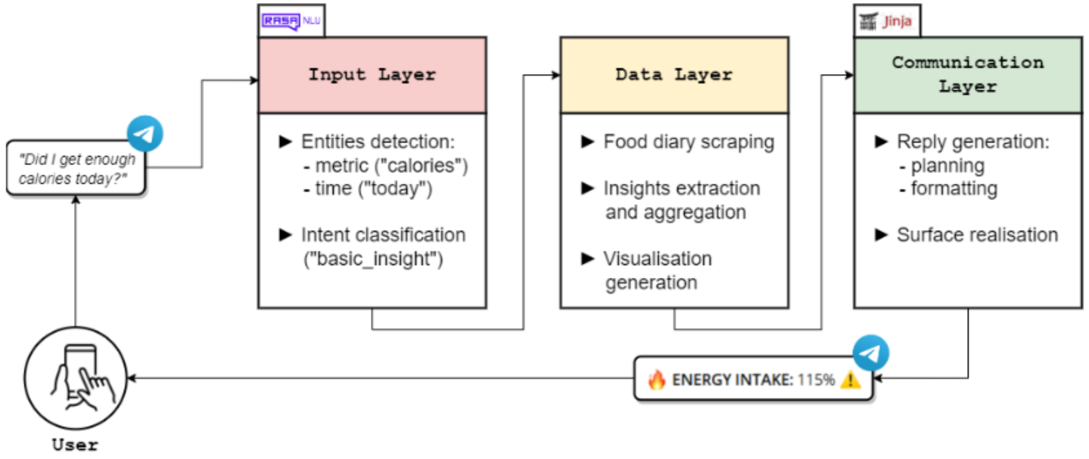


Figure 1: Chatbot architecture and interaction flow.

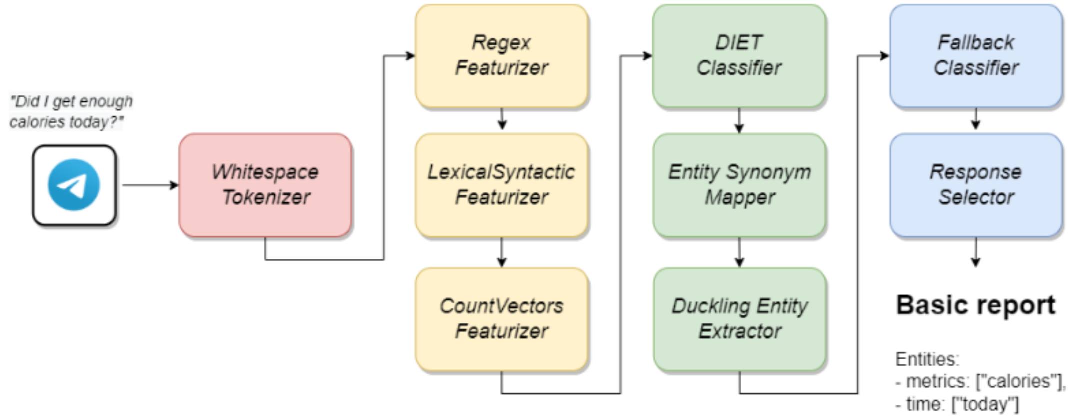


Figure 2: Overview of the NLU pipeline.

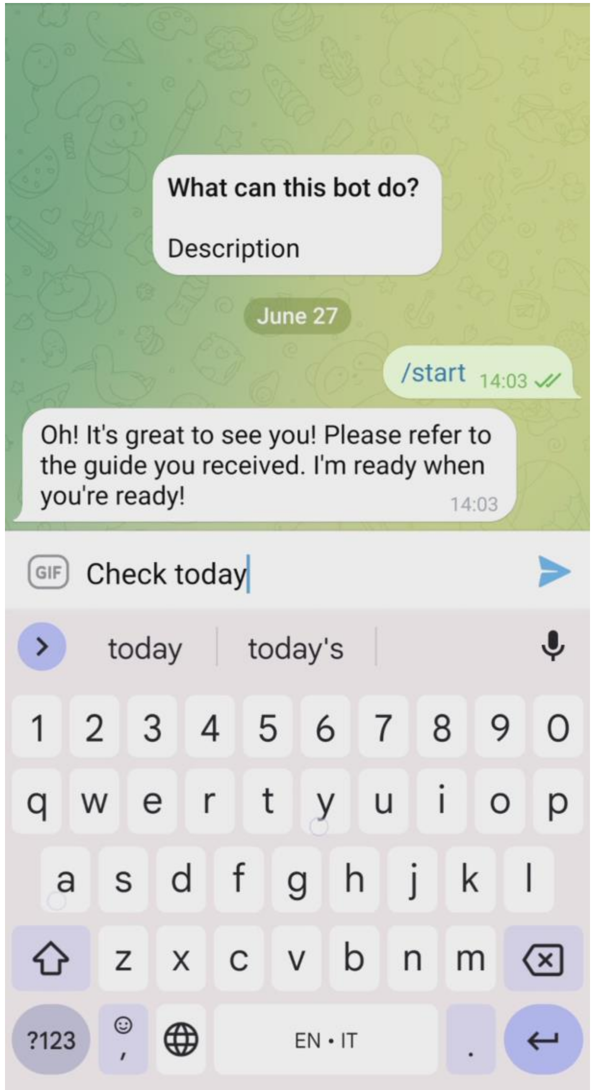
Overview of the chatbot

Overview of the chatbot

- Basic report: insights over a single time-frame

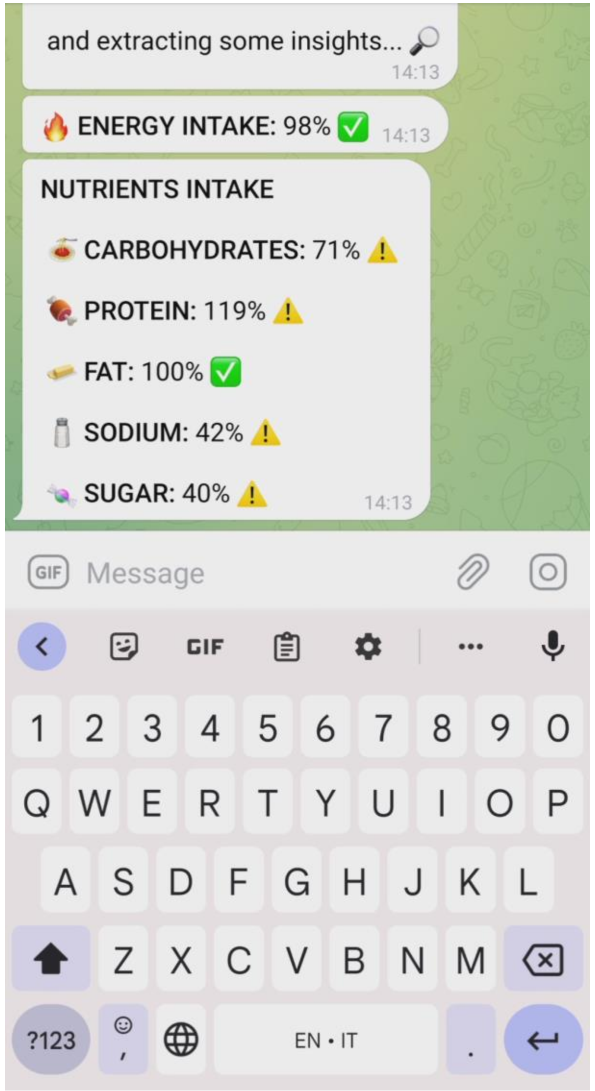
Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**



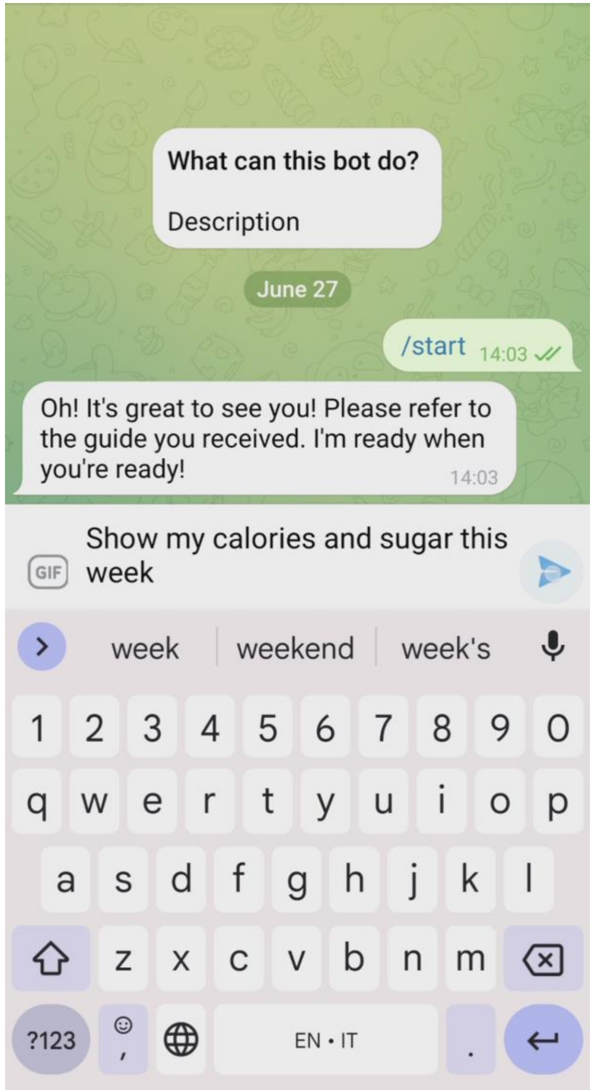
Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**



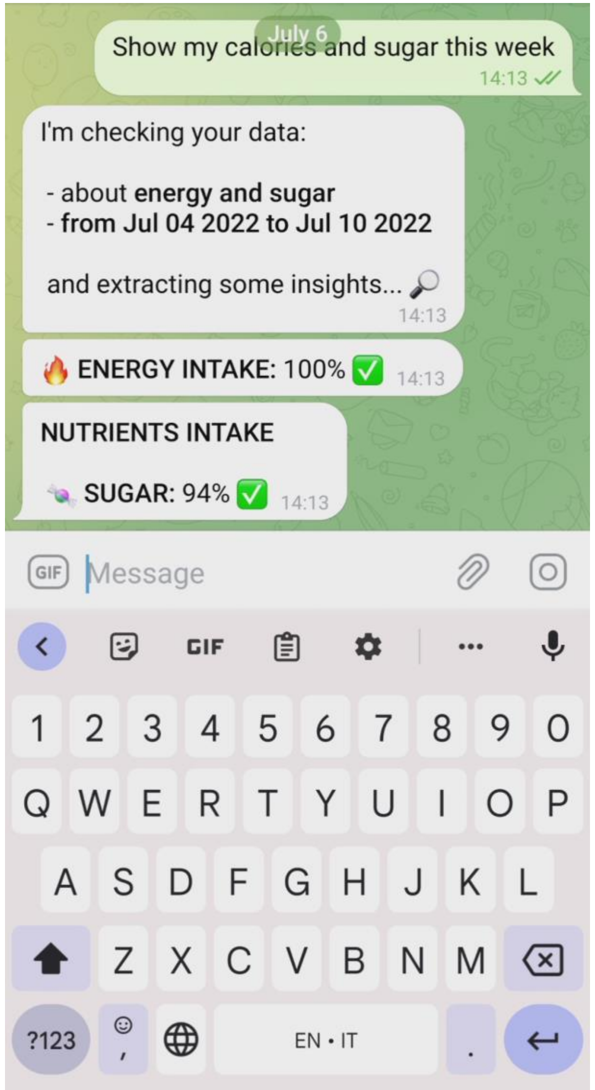
Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)



Overview of the chatbot

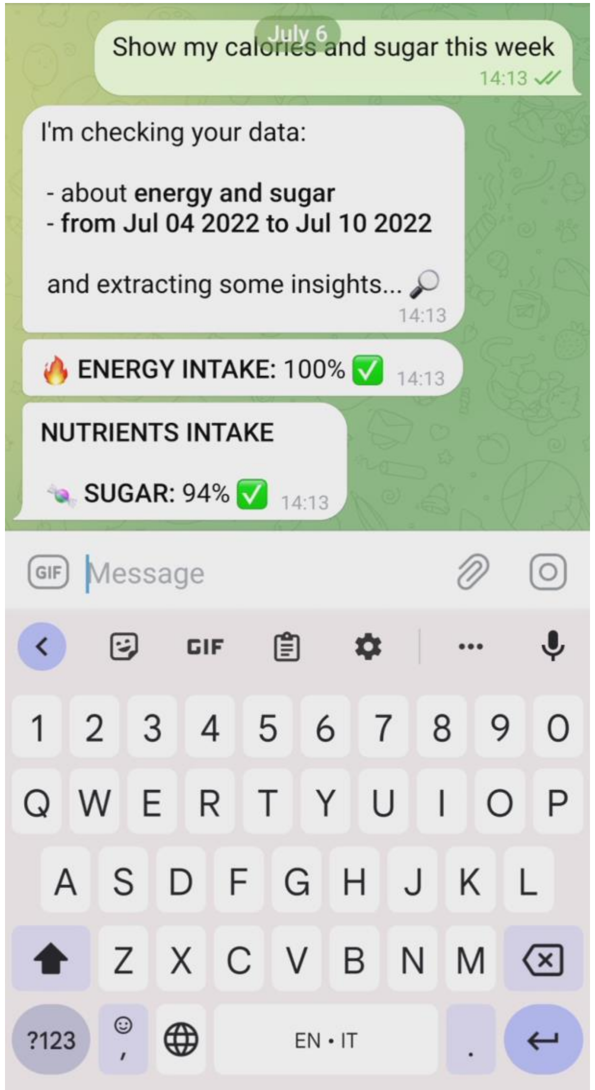
- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

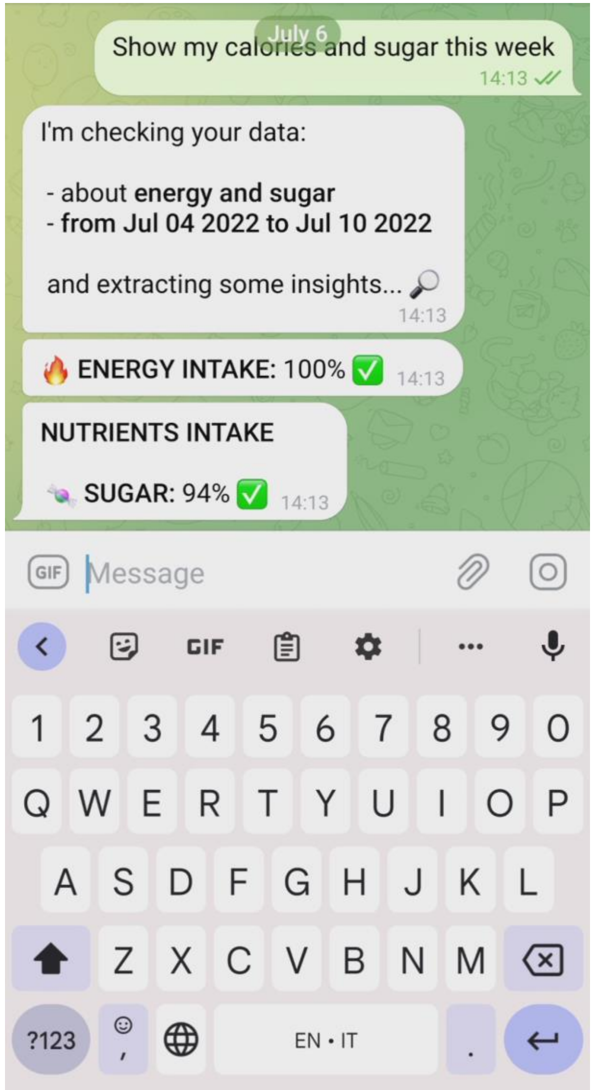
- Advanced insights:



Overview of the chatbot

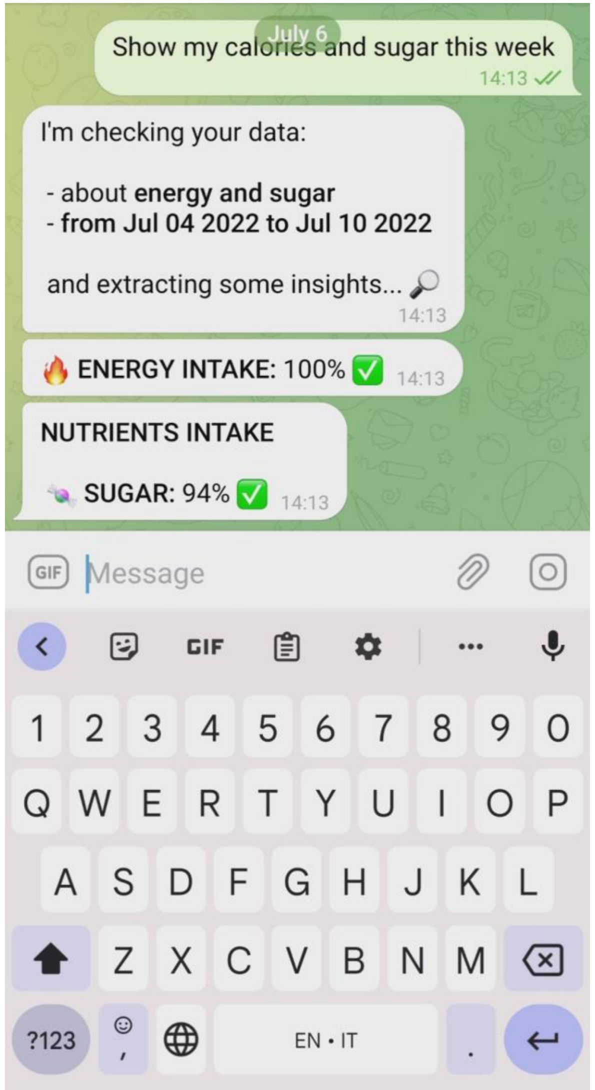
- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

- Advanced insights:
  - Combination of charts and texts



Overview of the chatbot

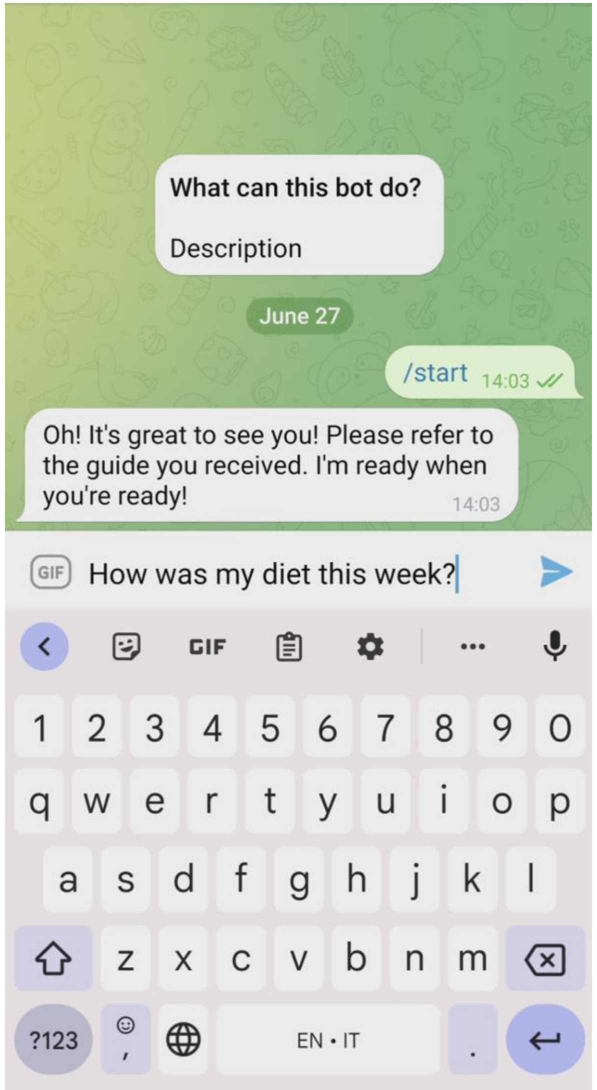
- **Basic report:** insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)
- **Advanced insights:**
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

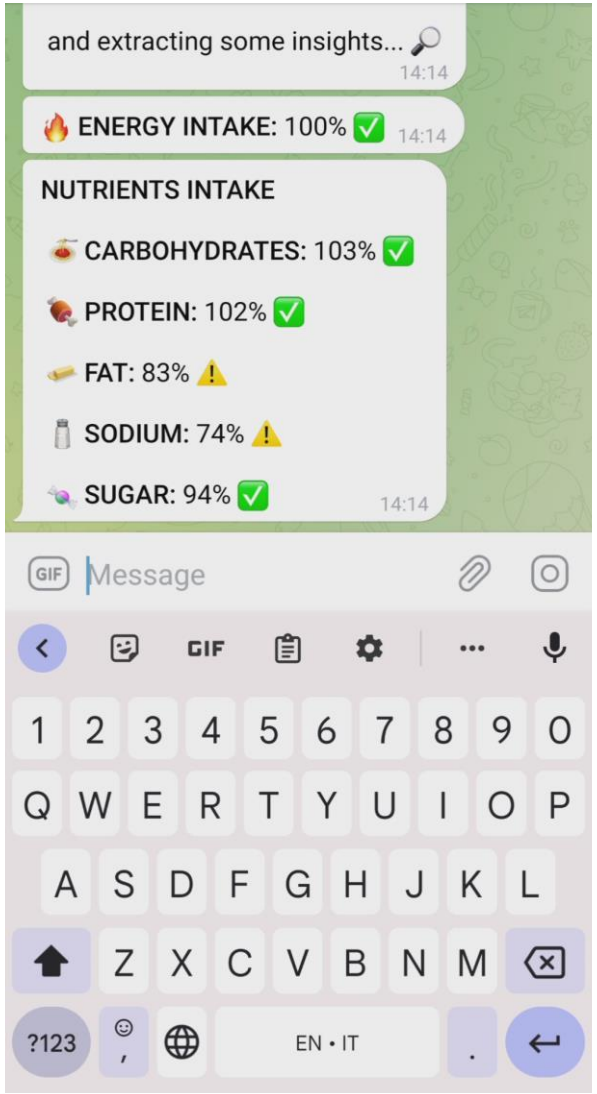
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

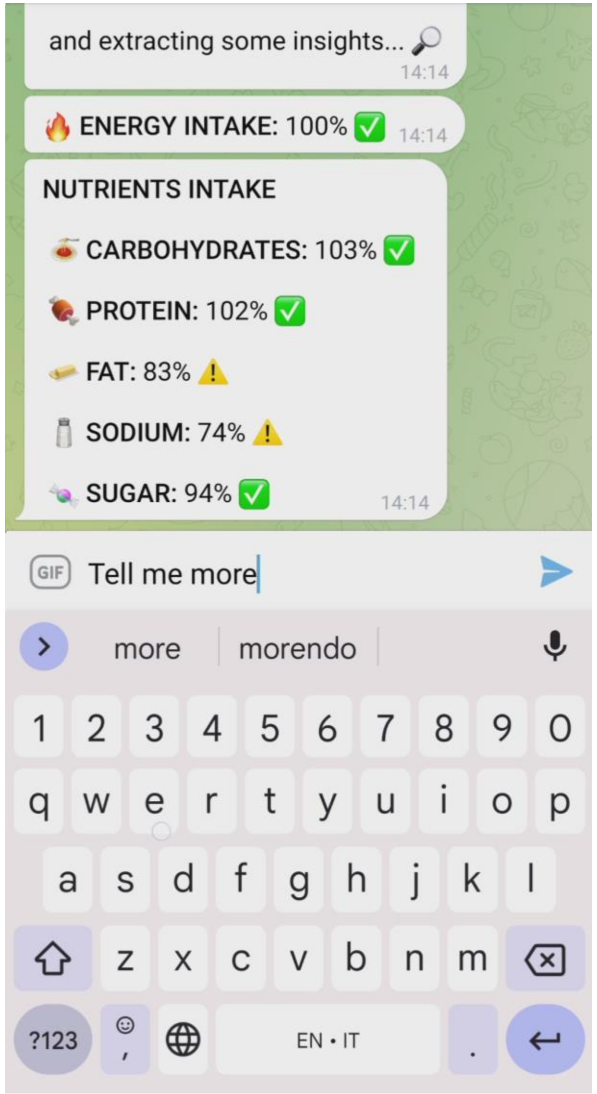
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

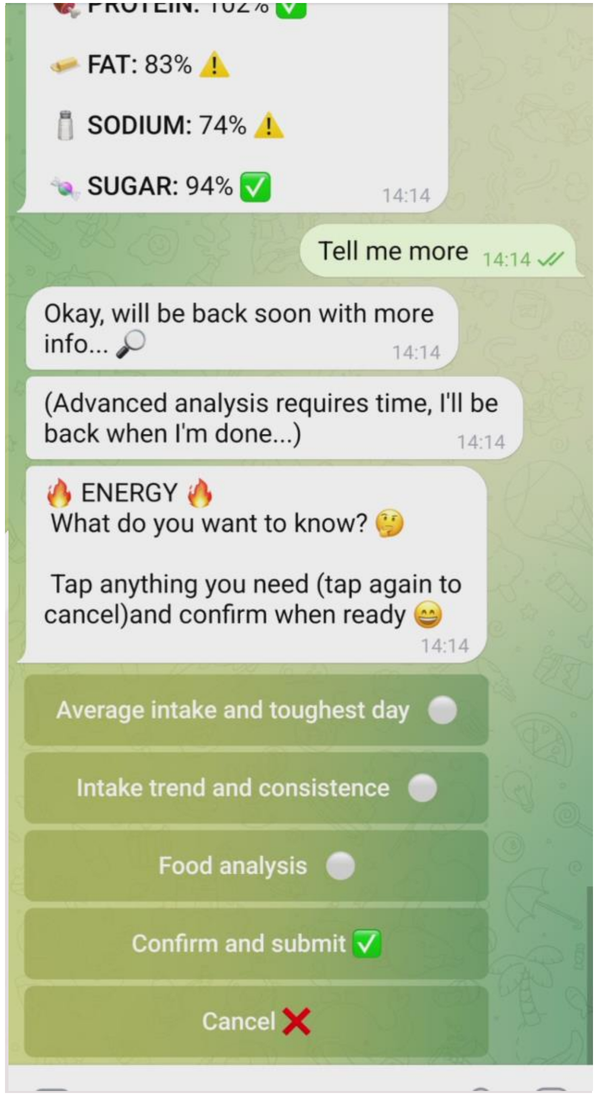
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

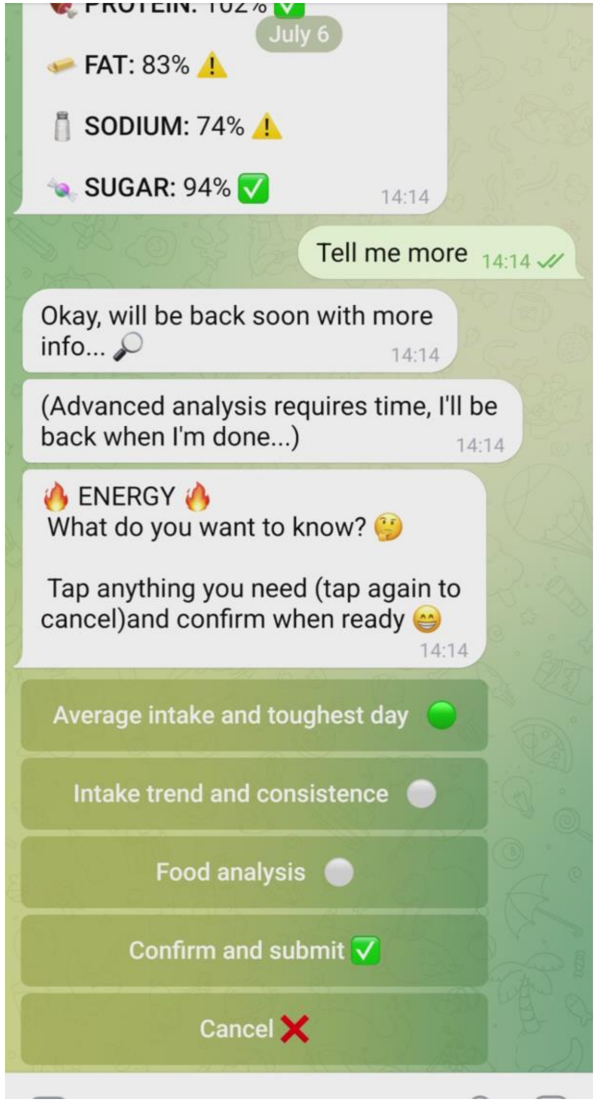
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights



Overview of the chatbot

- **Basic report:** insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

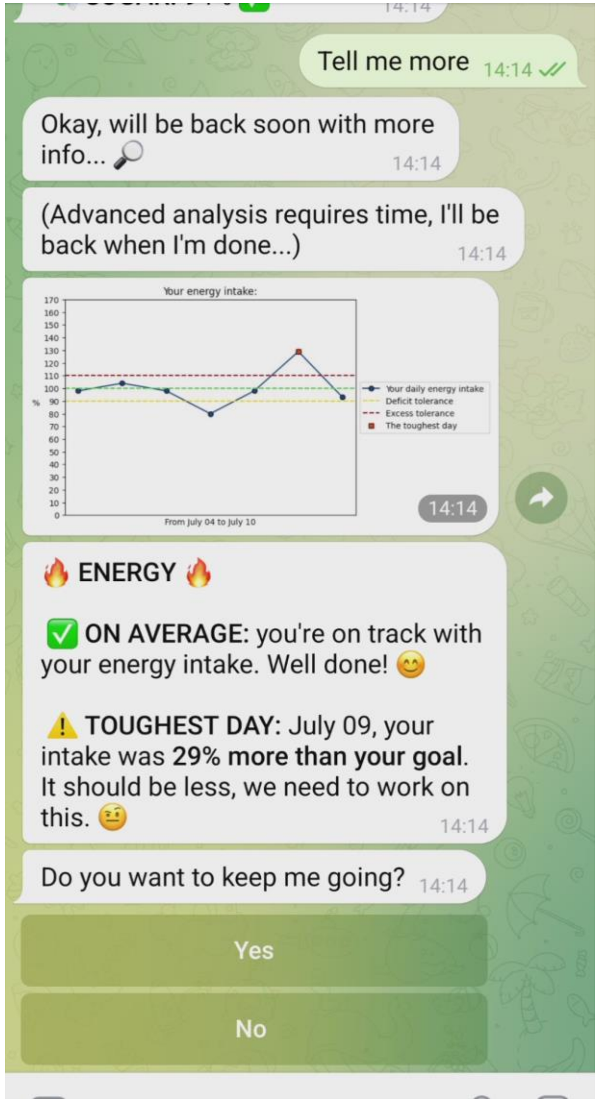
- **Advanced insights:**
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

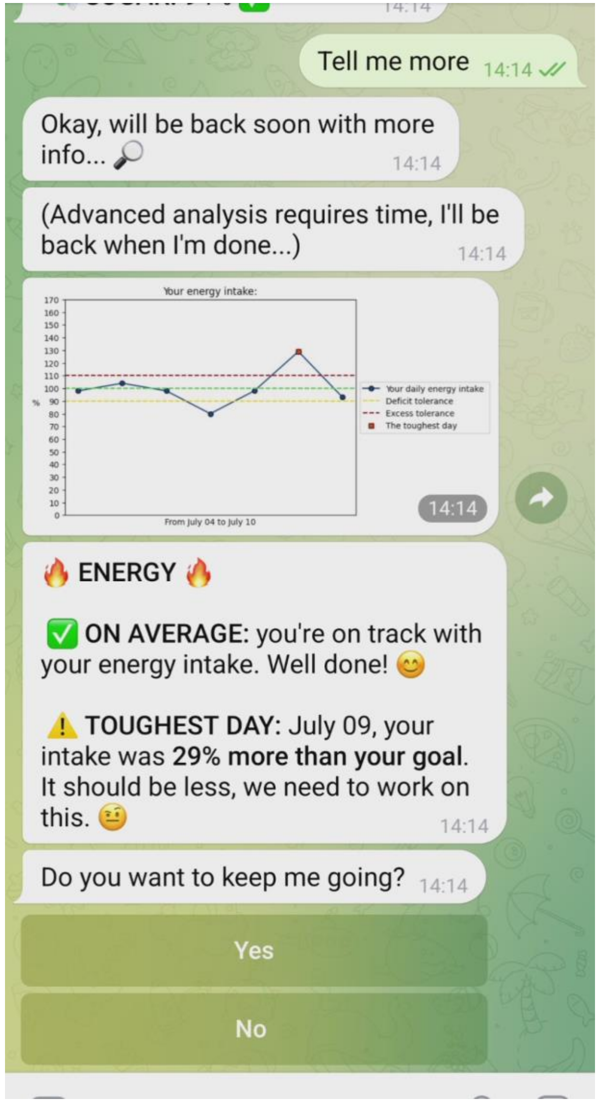
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

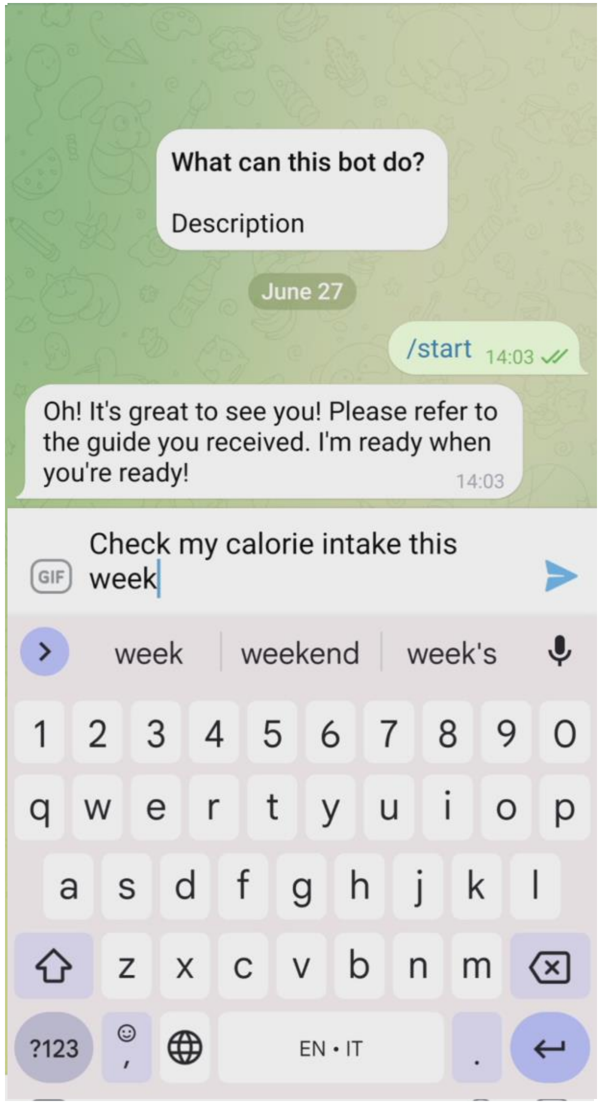
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily



Overview of the chatbot

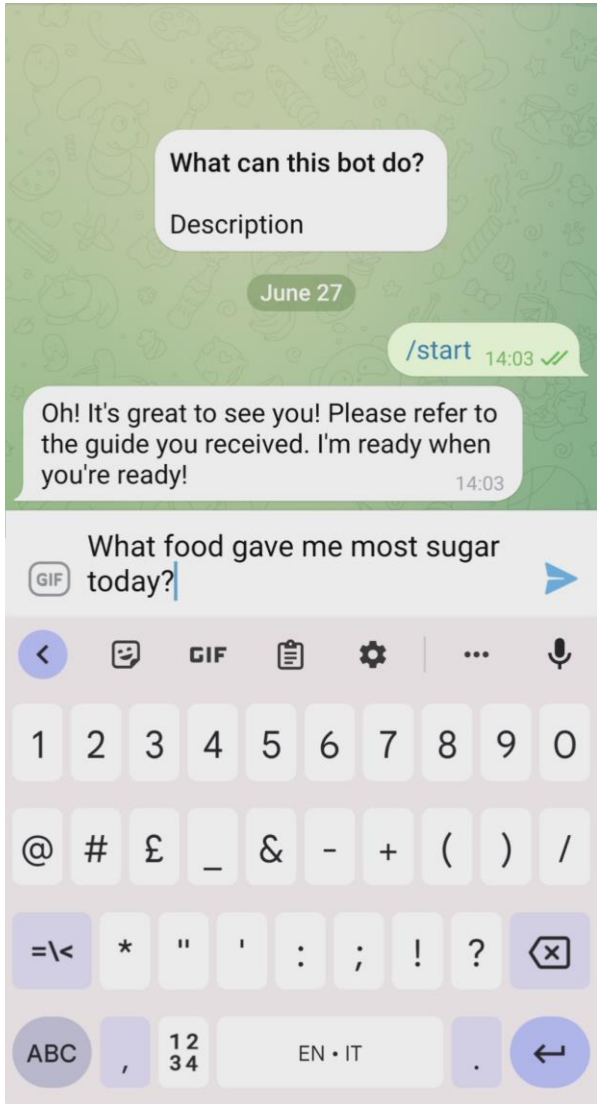
- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily

Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

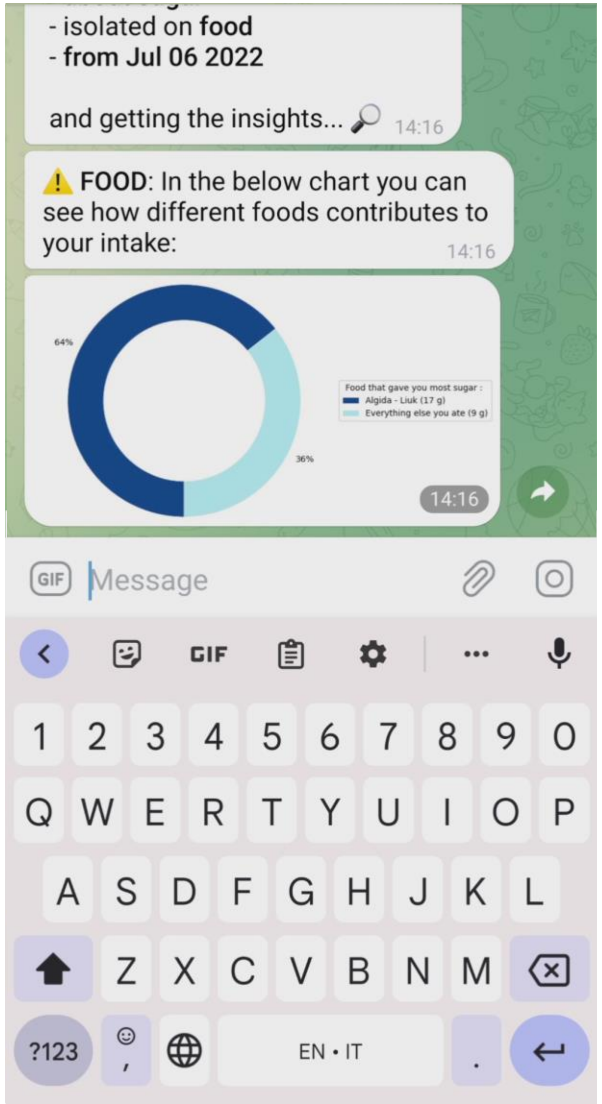
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily



Overview of the chatbot

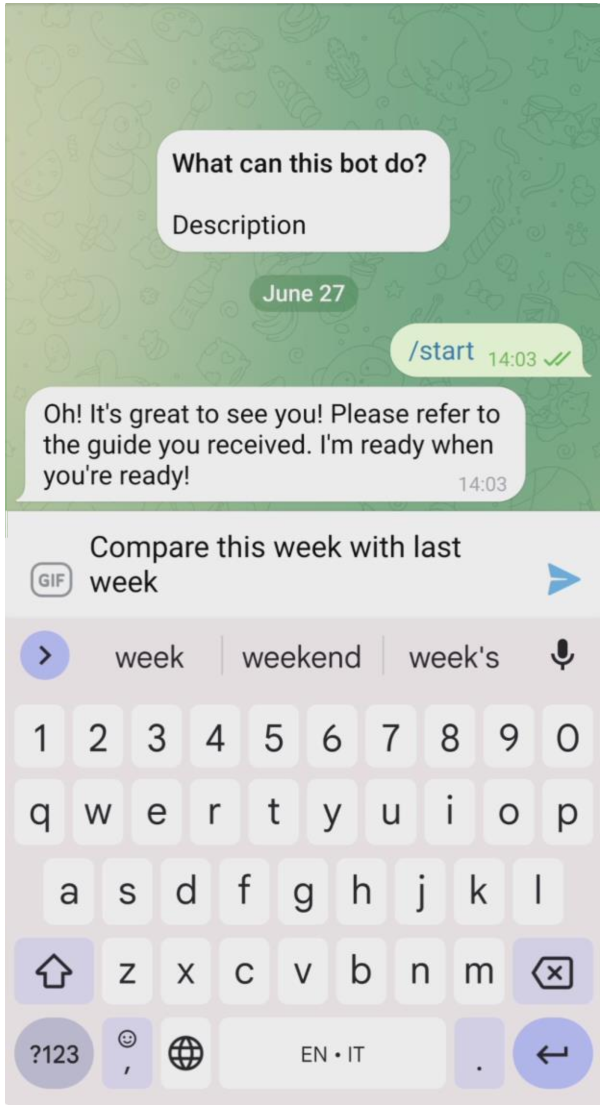
- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily
- Comparisons: insights over multiple time-frame

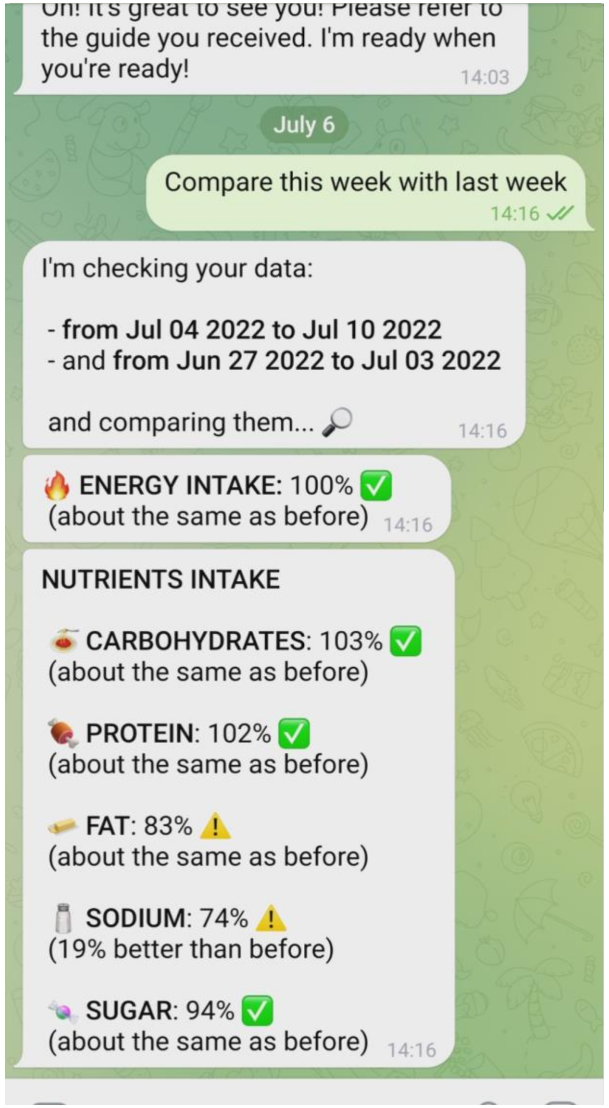


Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

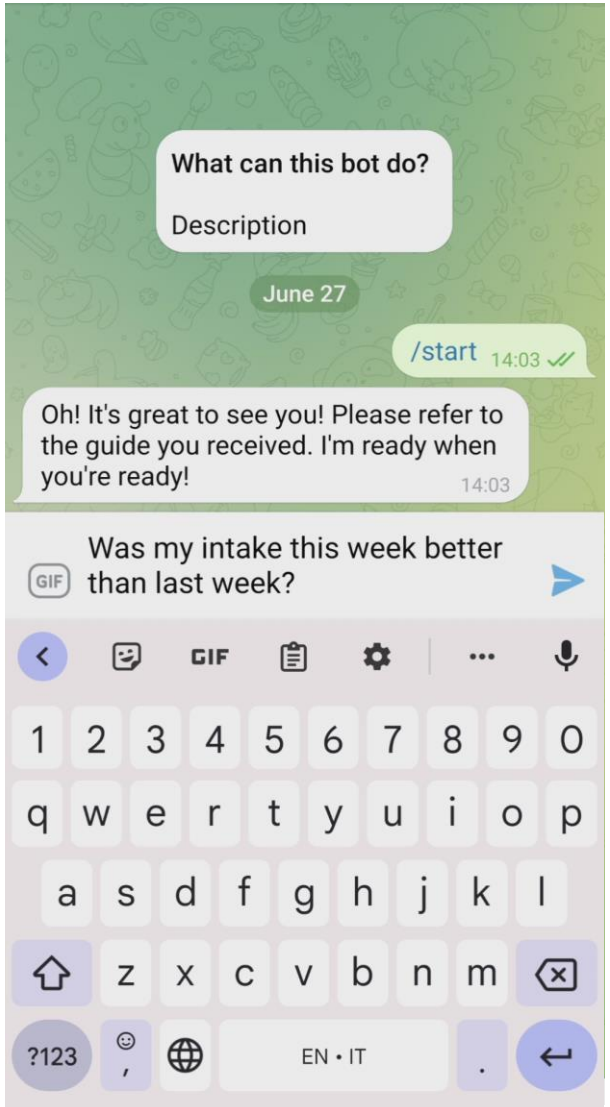
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily

- Comparisons: insights over multiple time-frame
  - **Extend** basic reports



Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)
- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily
- Comparisons: insights over multiple time-frame
  - **Extend** basic reports

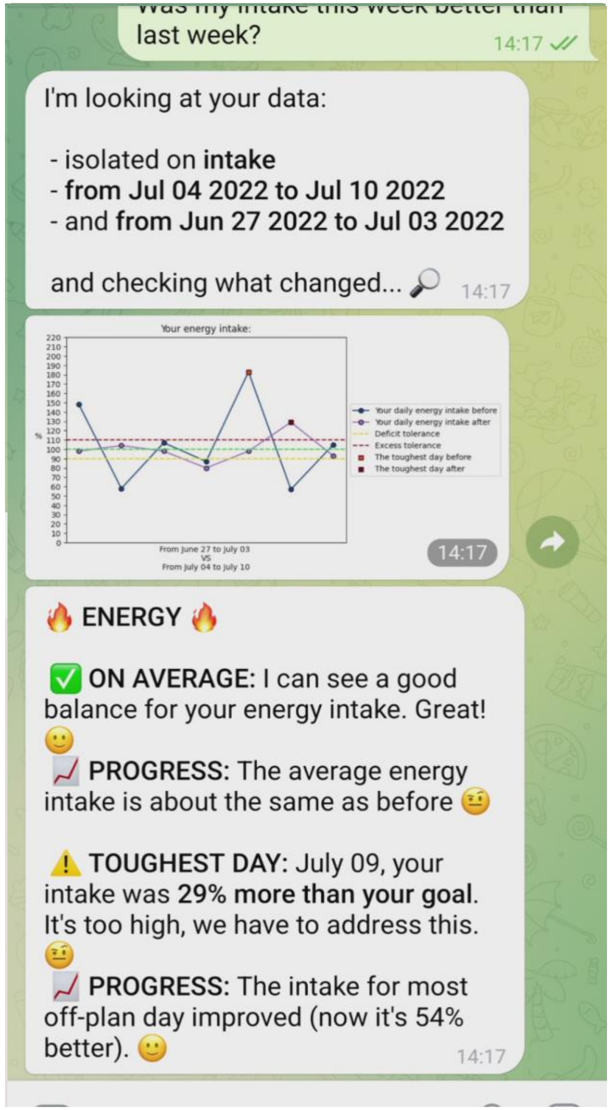


Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily

- Comparisons: insights over multiple time-frame
  - **Extend** basic reports

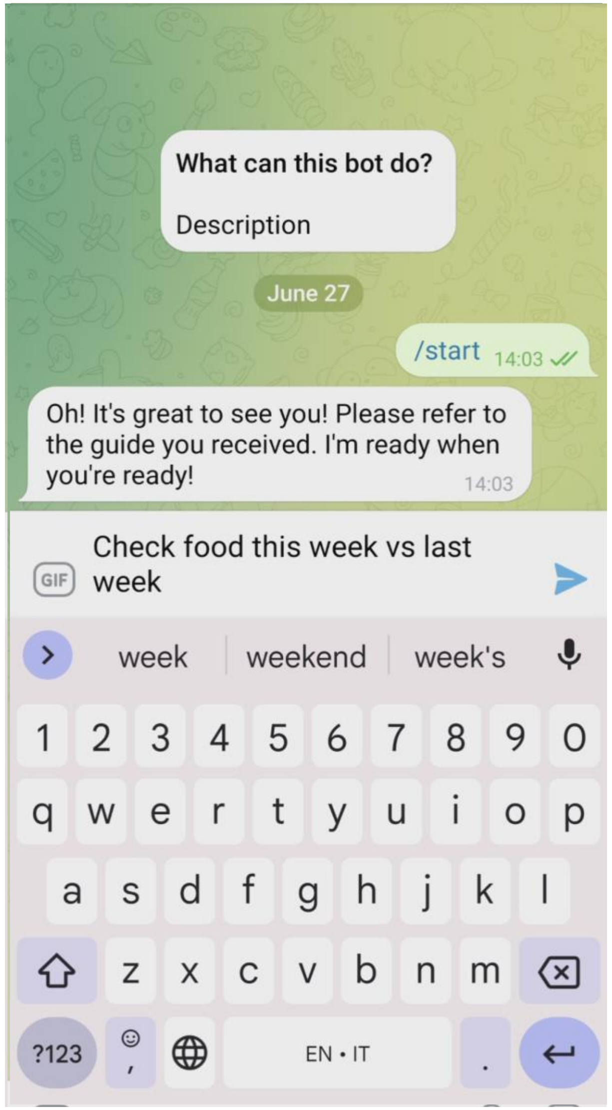


Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily

- Comparisons: insights over multiple time-frame
  - **Extend** basic reports

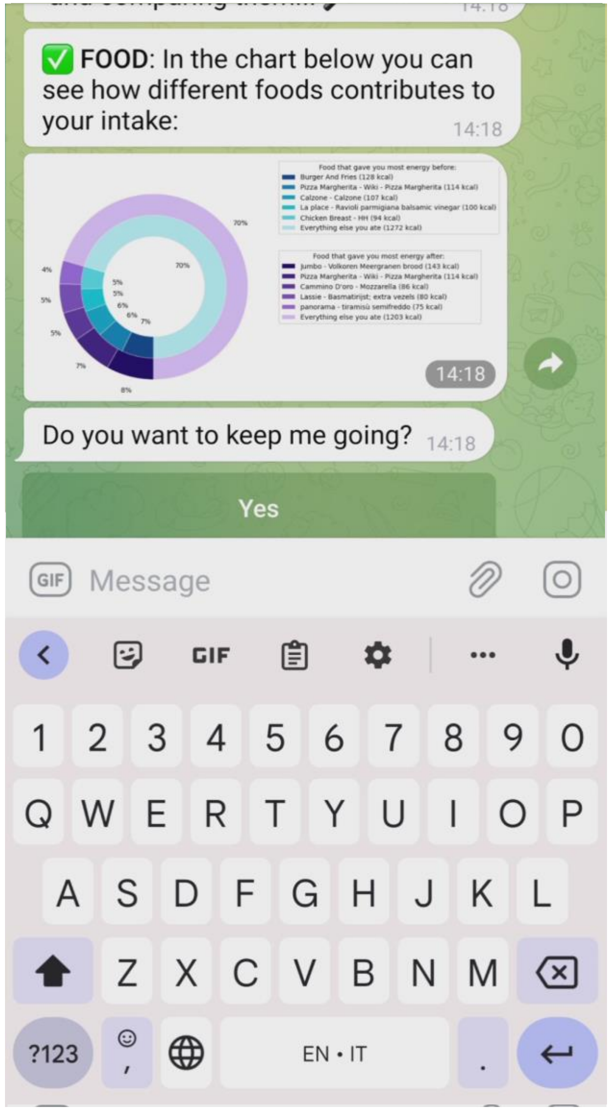


Overview of the chatbot

- Basic report: insights over a single time-frame
  - Starts with a **natural language query**
  - Users can specify **time** and **metrics** (calories/nutrients)

- Advanced insights:
  - Combination of charts and texts
  - **Guided navigation:** for inexperienced users
    - Starts from a basic report
    - **Button interface** with available insights
  - **Natural language query:** for advanced users
    - Triggered arbitrarily

- Comparisons: insights over multiple time-frame
  - **Extend** basic reports



Experiment setup

Experiment setup

- 87 participants (from Amazon Mechanical Turk)

Experiment setup

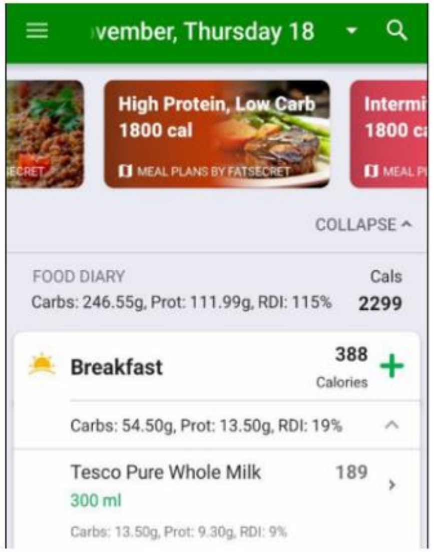
- 87 participants (from Amazon Mechanical Turk)
    - Each worker was randomly assigned to a tool

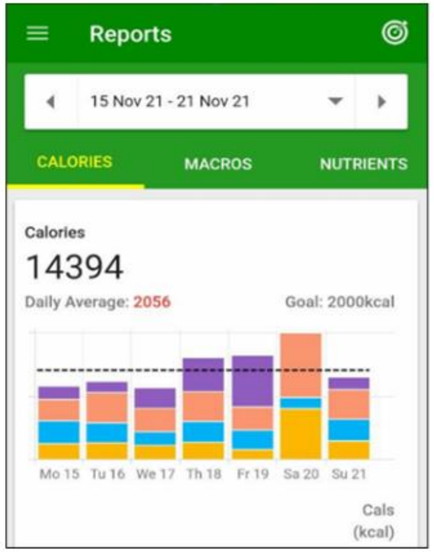
Experiment setup

- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)

Experiment setup

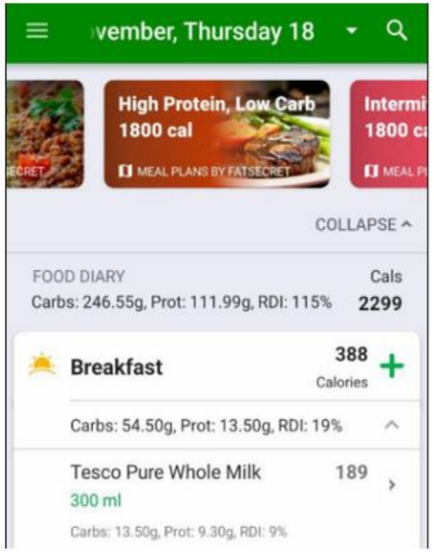
- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)

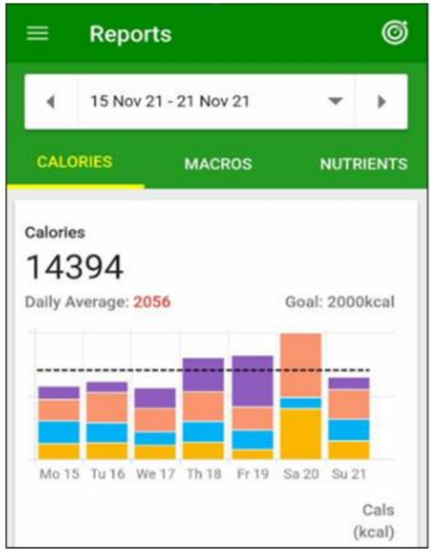


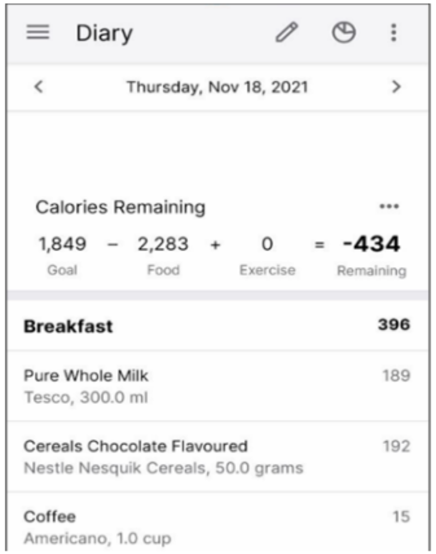


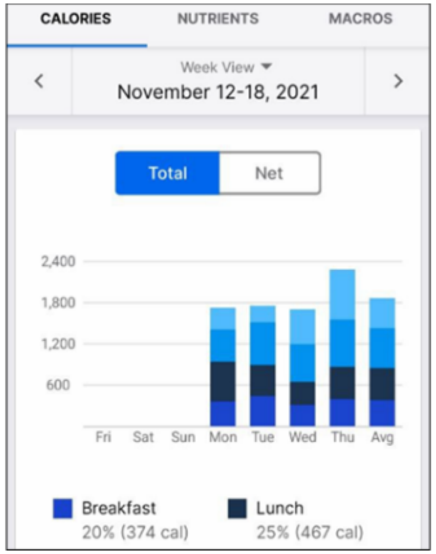
Experiment setup

- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)



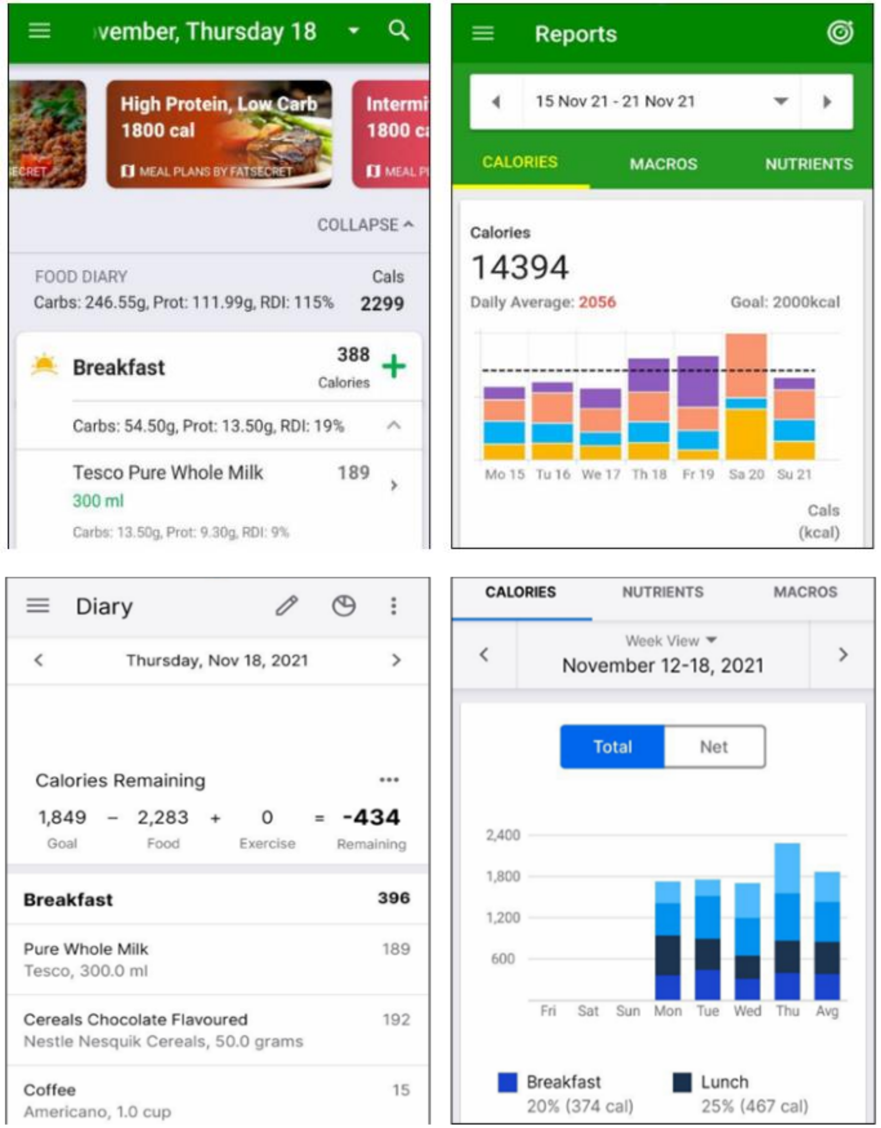






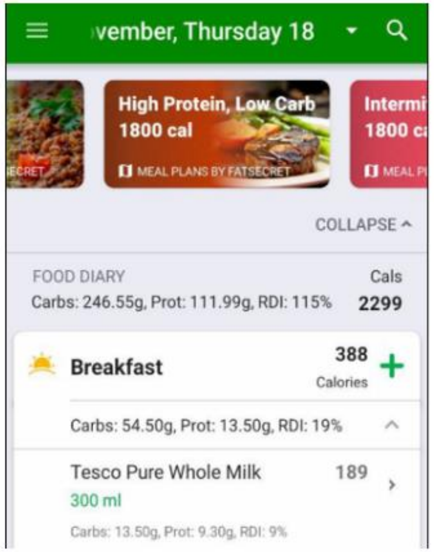
Experiment setup

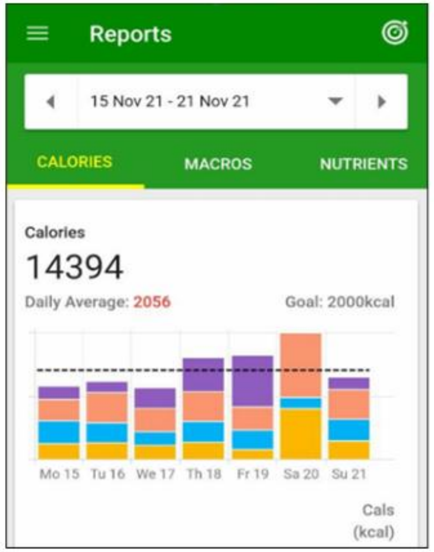
- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)
  - We fill every tool with a **simulated food diary** (2 weeks of meals)

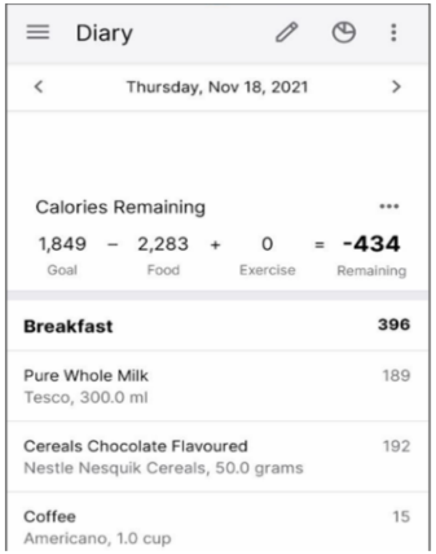


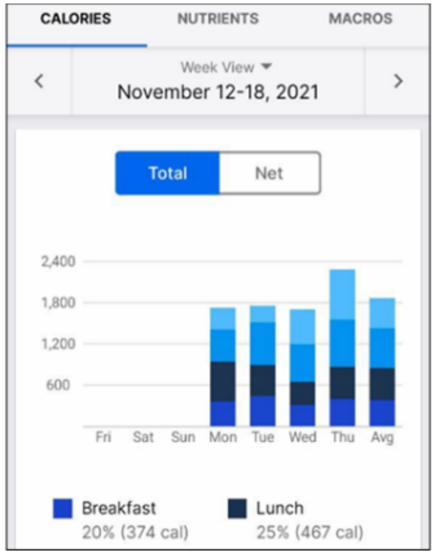
Experiment setup

- 87 participants (from **Amazon Mechanical Turk**)
    - Each worker was randomly assigned to a tool
        - **Our chatbot** (CB)
        - FatSecret (FS)
        - MyFitnessPal (MFP)
    - We fill every tool with a **simulated food diary** (2 weeks of meals)
    - We give every worker:



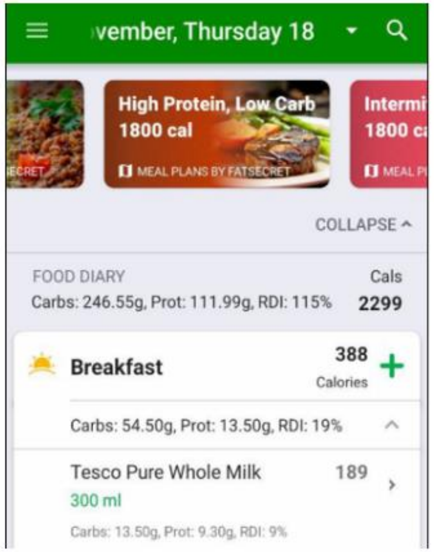


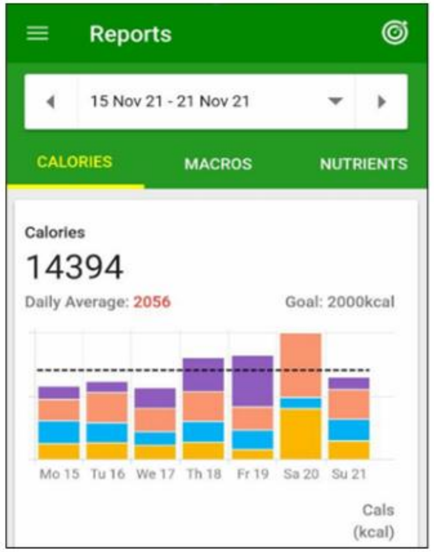


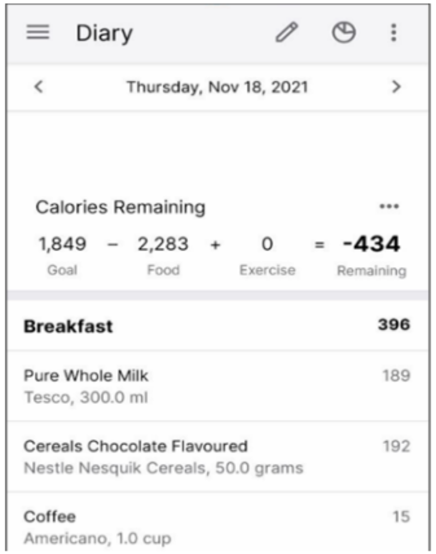


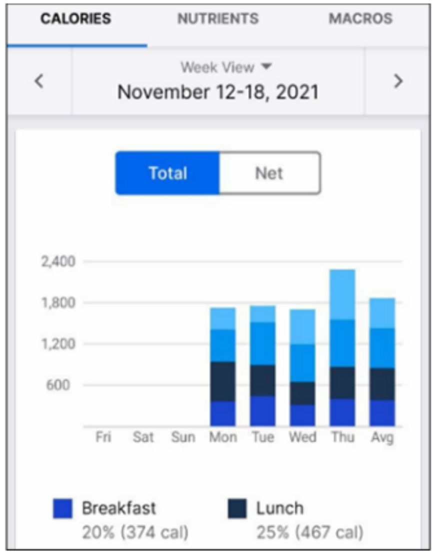
Experiment setup

- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)
  - We fill every tool with a **simulated food diary** (2 weeks of meals)
  - We give every worker:
    - A **guide** for the tool



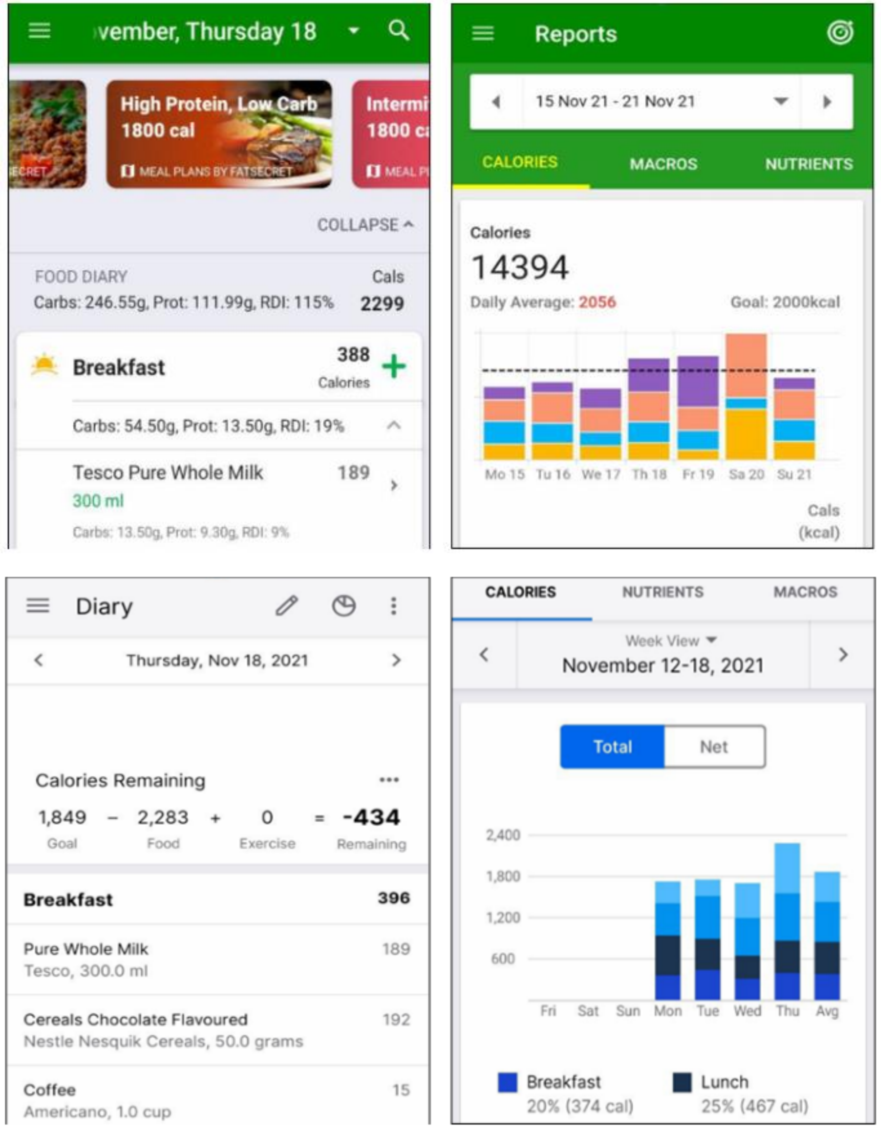






Experiment setup

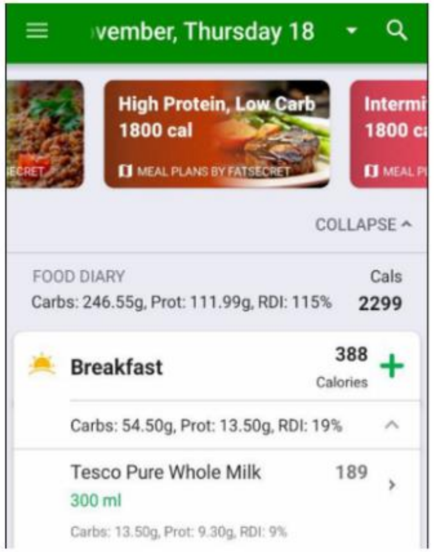
- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)
  - We fill every tool with a **simulated food diary** (2 weeks of meals)
  - We give every worker:
    - A **guide** for the tool
    - A **glossary** of experiment terminology

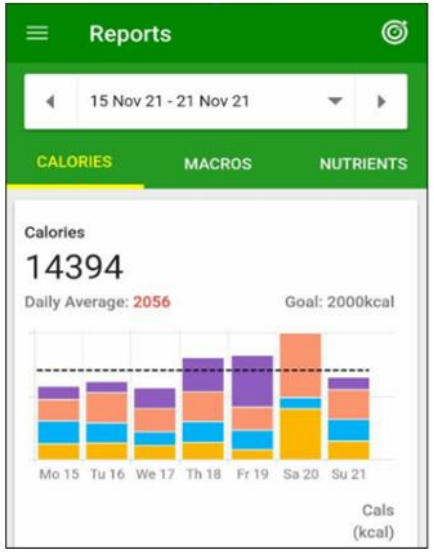


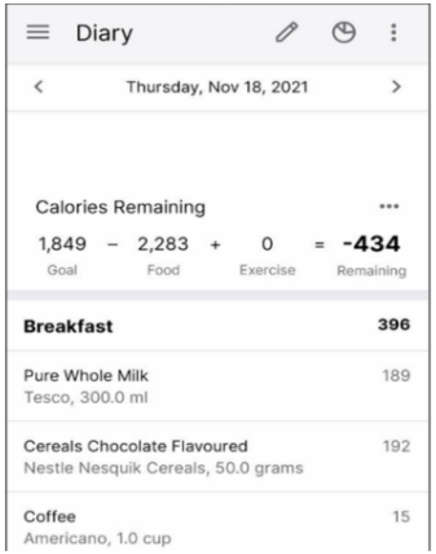
Experiment setup

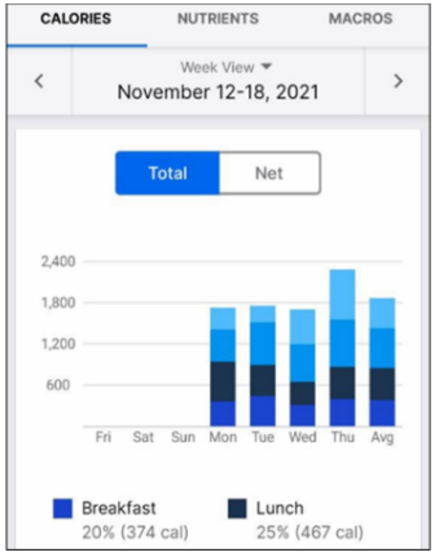
- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)
  - We fill every tool with a **simulated food diary** (2 weeks of meals)
  - We give every worker:
    - A **guide** for the tool
    - A **glossary** of experiment terminology

- We evaluate:





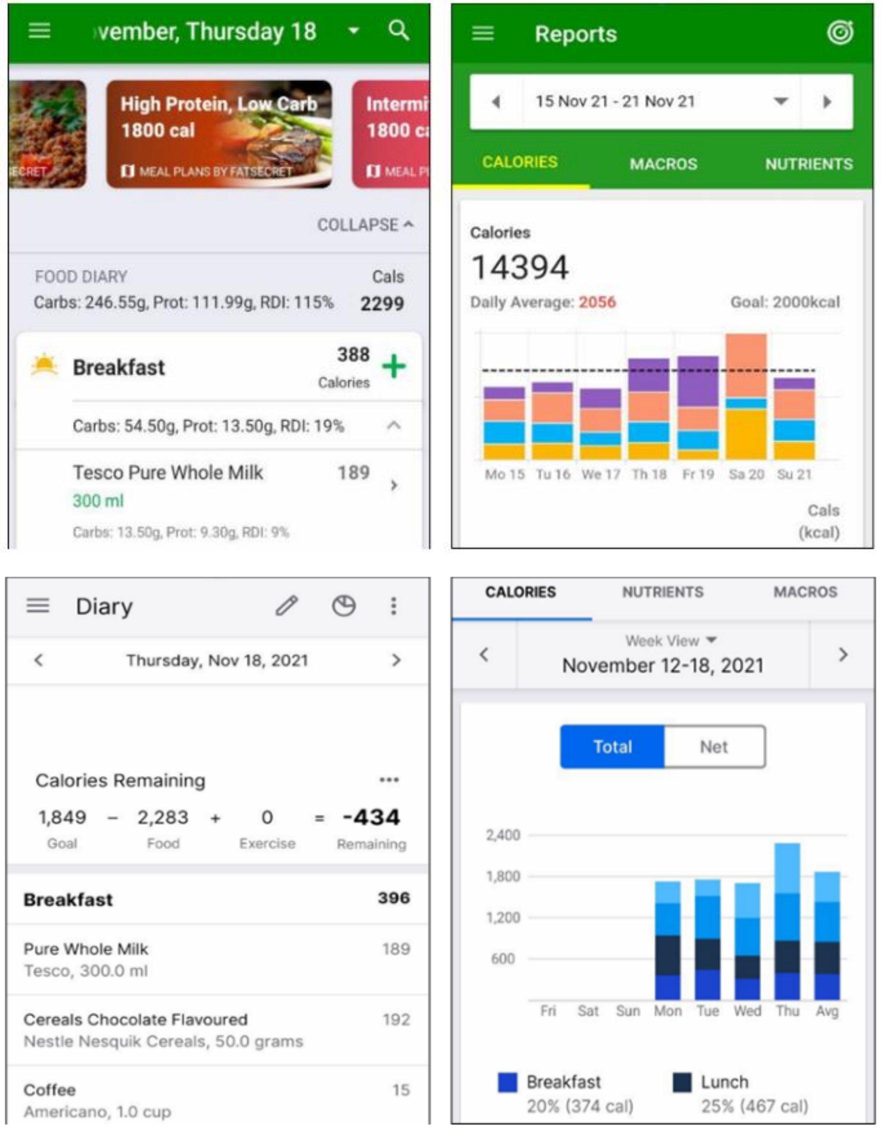




Experiment setup

- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)
  - We fill every tool with a **simulated food diary** (2 weeks of meals)
  - We give every worker:
    - A **guide** for the tool
    - A **glossary** of experiment terminology

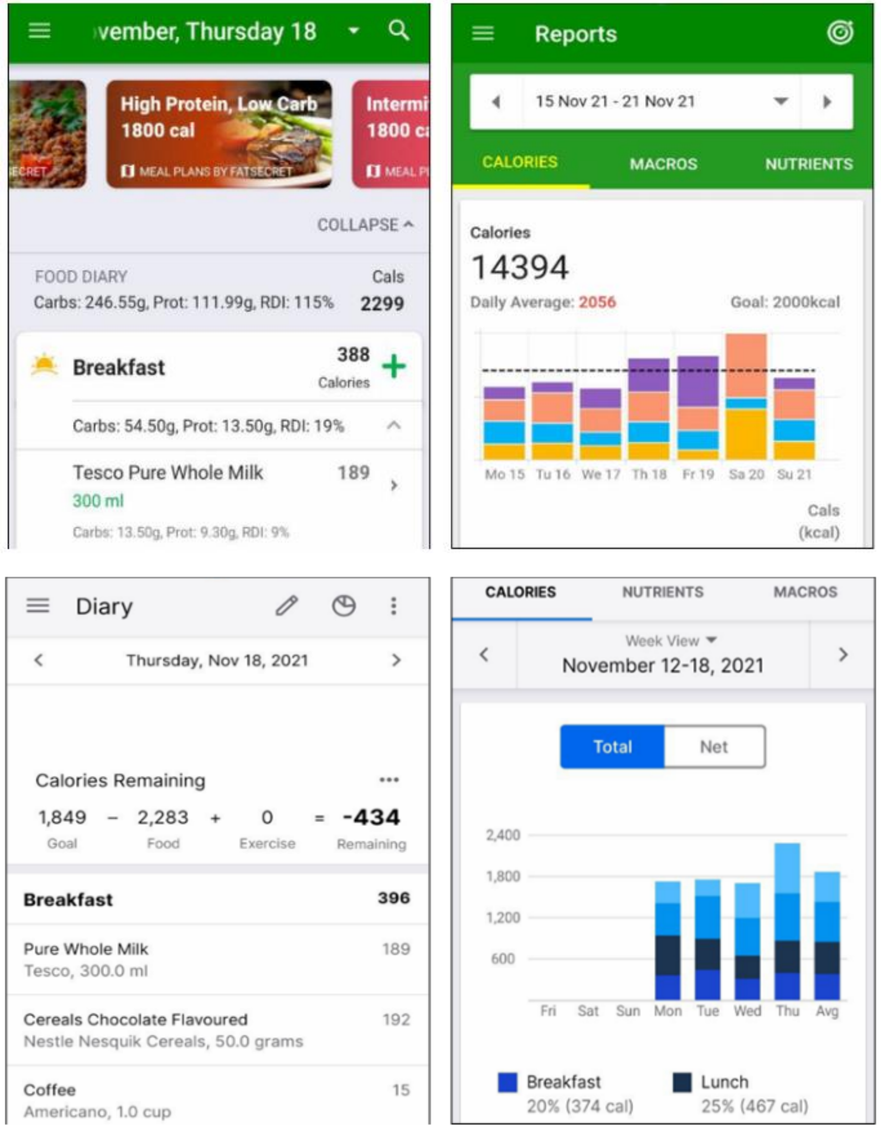
- We evaluate:
  - **Informativeness** of each tool



Experiment setup

- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)
  - We fill every tool with a **simulated food diary** (2 weeks of meals)
  - We give every worker:
    - A **guide** for the tool
    - A **glossary** of experiment terminology

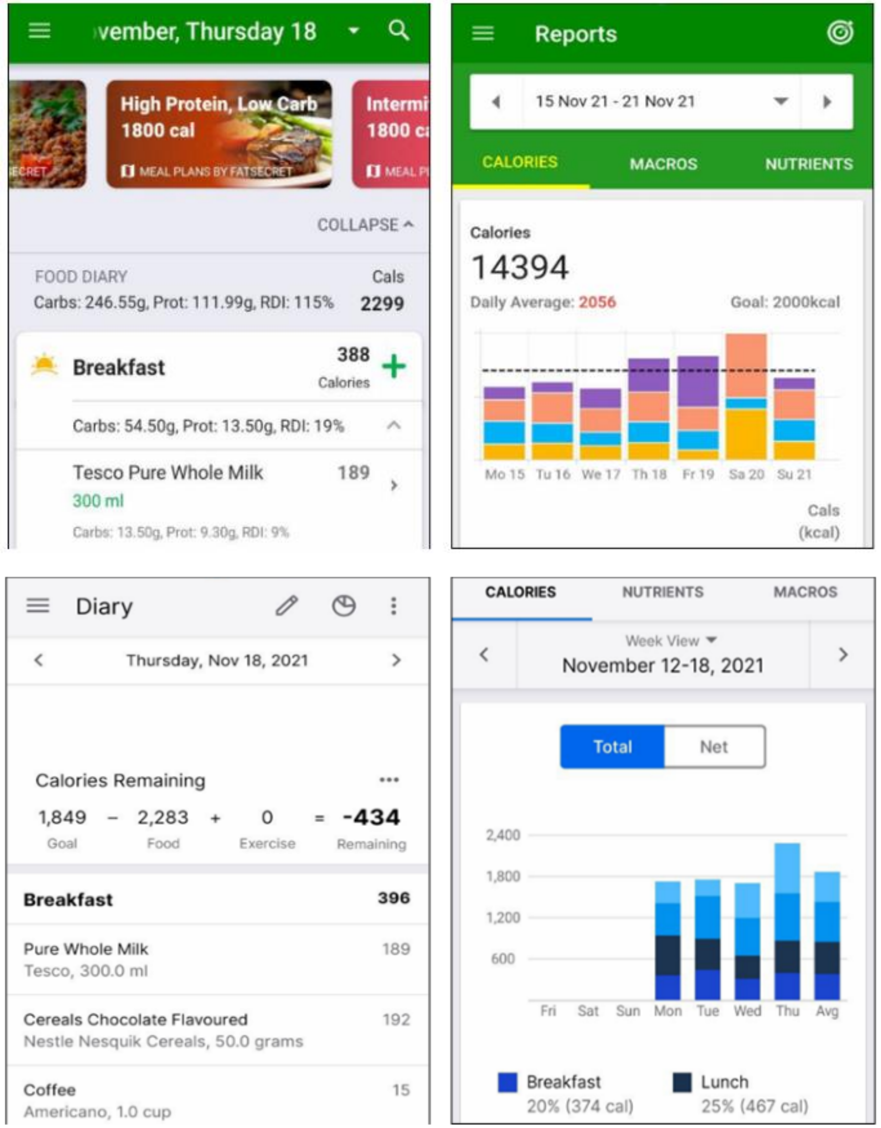
- We evaluate:
  - **Informativeness** of each tool
  - Inter-sample **nutrition literacy**



Experiment setup

- 87 participants (from **Amazon Mechanical Turk**)
  - Each worker was randomly assigned to a tool
    - **Our chatbot** (CB)
    - FatSecret (FS)
    - MyFitnessPal (MFP)
  - We fill every tool with a **simulated food diary** (2 weeks of meals)
  - We give every worker:
    - A **guide** for the tool
    - A **glossary** of experiment terminology

- We evaluate:
  - **Informativeness** of each tool
  - Inter-sample **nutrition literacy**
  - Workers' **perception of the tool**



Measuring informativeness

Measuring informativeness

- “How successfully a person is able to convey an intended message” (Webster and Morris, 2019):

Measuring informativeness

- "How successfully a ~~person~~ tool is able to convey an intended message" (Webster and Morris, 2019):

Measuring informativeness

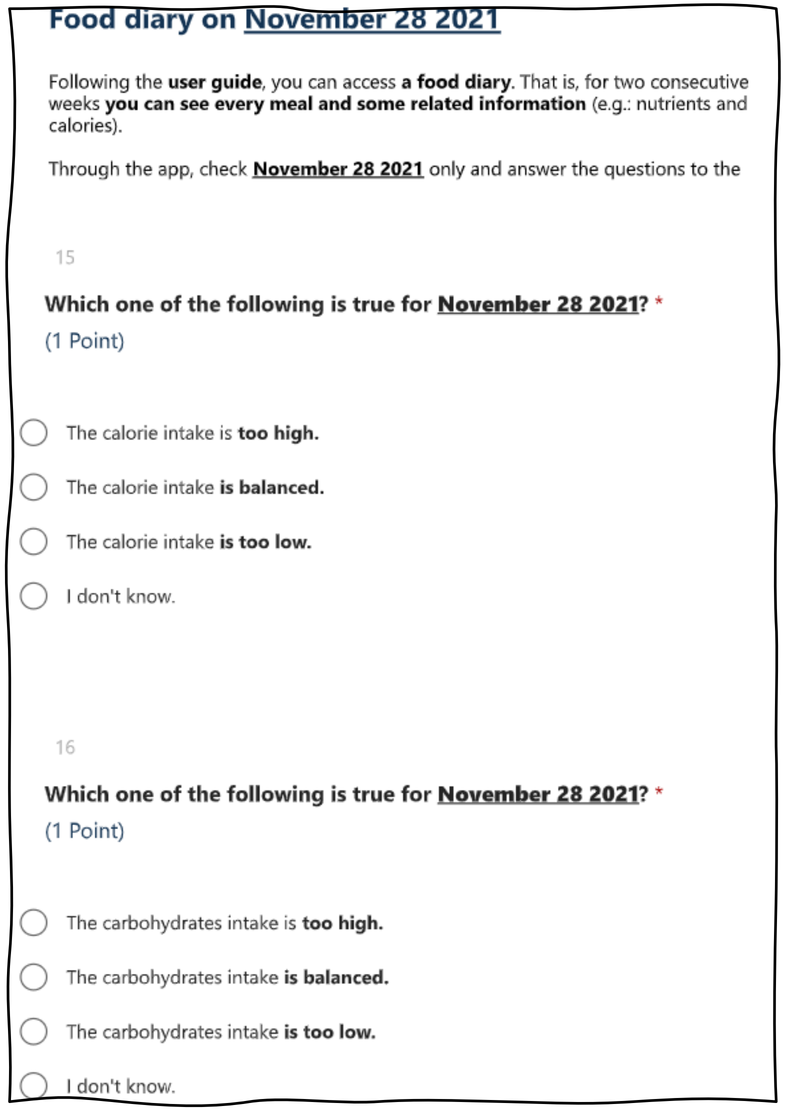
- “How successfully a ~~person~~ tool is able to convey an intended message” (Webster and Morris, 2019):

- Workers used their assigned tool to take a 10-point quiz

Measuring informativeness

- "How successfully a ~~person~~ tool is able to convey an intended message" (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day



Food diary on November 28 2021

Following the **user guide**, you can access **a food diary**. That is, for two consecutive weeks **you can see every meal and some related information** (e.g.: nutrients and calories).

Through the app, check **November 28 2021** only and answer the questions to the

15

Which one of the following is true for November 28 2021? *

(1 Point)

- The calorie intake is **too high.**
- The calorie intake **is balanced.**
- The calorie intake **is too low.**
- I don't know.

16

Which one of the following is true for November 28 2021? *

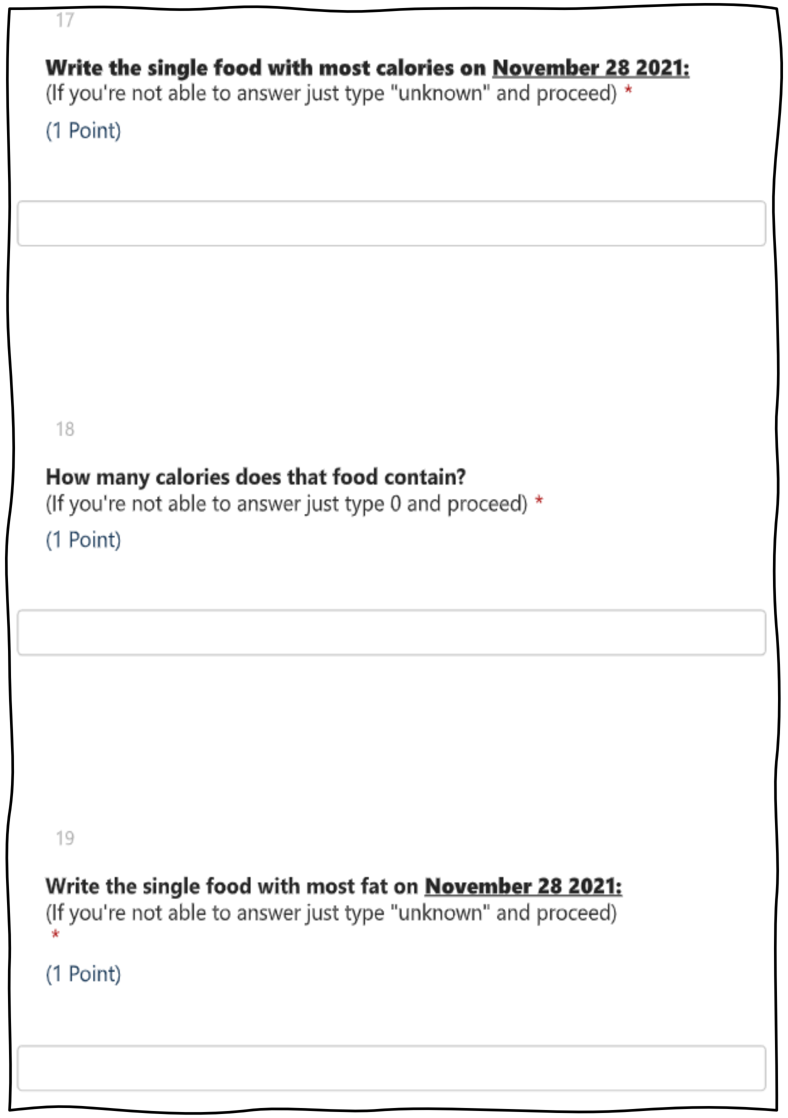
(1 Point)

- The carbohydrates intake is **too high.**
- The carbohydrates intake **is balanced.**
- The carbohydrates intake **is too low.**
- I don't know.

Measuring informativeness

- “How successfully a ~~person~~ tool is able to convey an intended message” (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet



17

Write the single food with most calories on <u>November 28 2021:</u>
(If you're not able to answer just type "unknown" and proceed) *

(1 Point)

18

How many calories does that food contain?
(If you're not able to answer just type 0 and proceed) *

(1 Point)

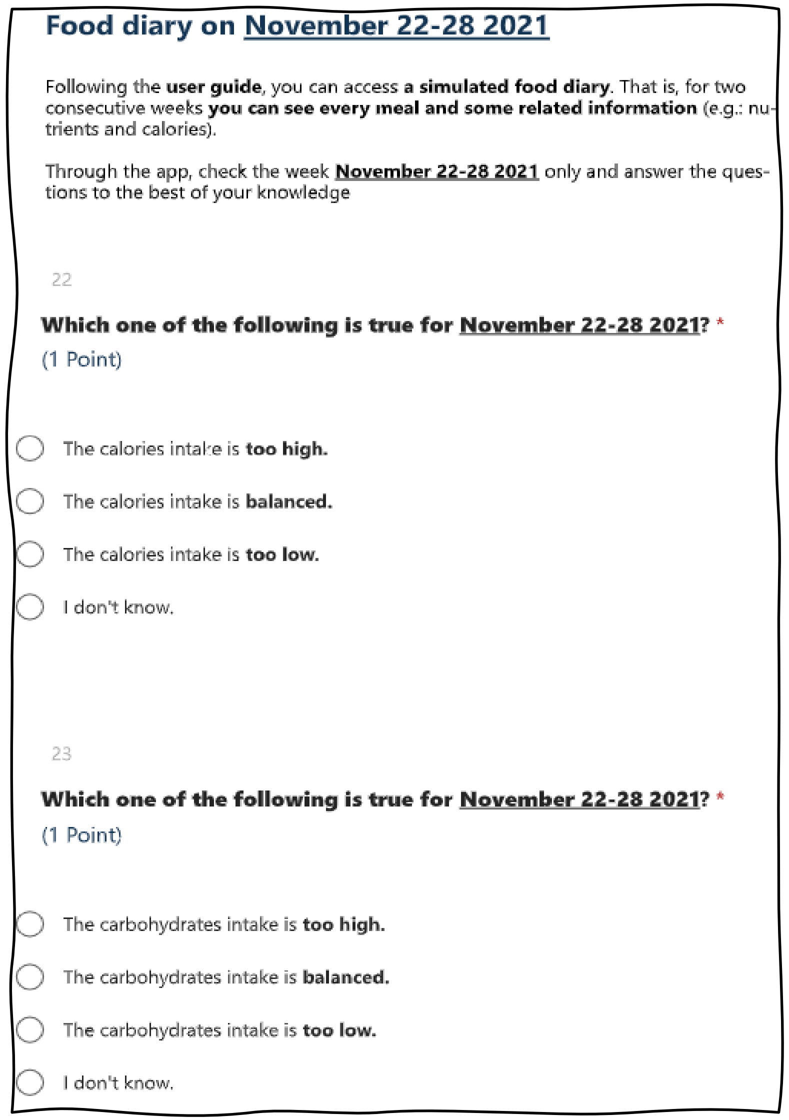
19

Write the single food with most fat on <u>November 28 2021:</u>
(If you're not able to answer just type "unknown" and proceed)
*

(1 Point)

Measuring informativeness

- “How successfully a ~~person~~ tool is able to convey an intended message” (**Webster and Morris, 2019**):
- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet
  - **Week analysis (2pt):** intake balance across a week



Food diary on November 22-28 2021

Following the **user guide**, you can access **a simulated food diary**. That is, for two consecutive weeks **you can see every meal and some related information** (e.g.: nutrients and calories).

Through the app, check the week **November 22-28 2021** only and answer the questions to the best of your knowledge

22

Which one of the following is true for November 22-28 2021? *

(1 Point)

- The calories intake is **too high.**
- The calories intake is **balanced.**
- The calories intake is **too low.**
- I don't know.

23

Which one of the following is true for November 22-28 2021? *

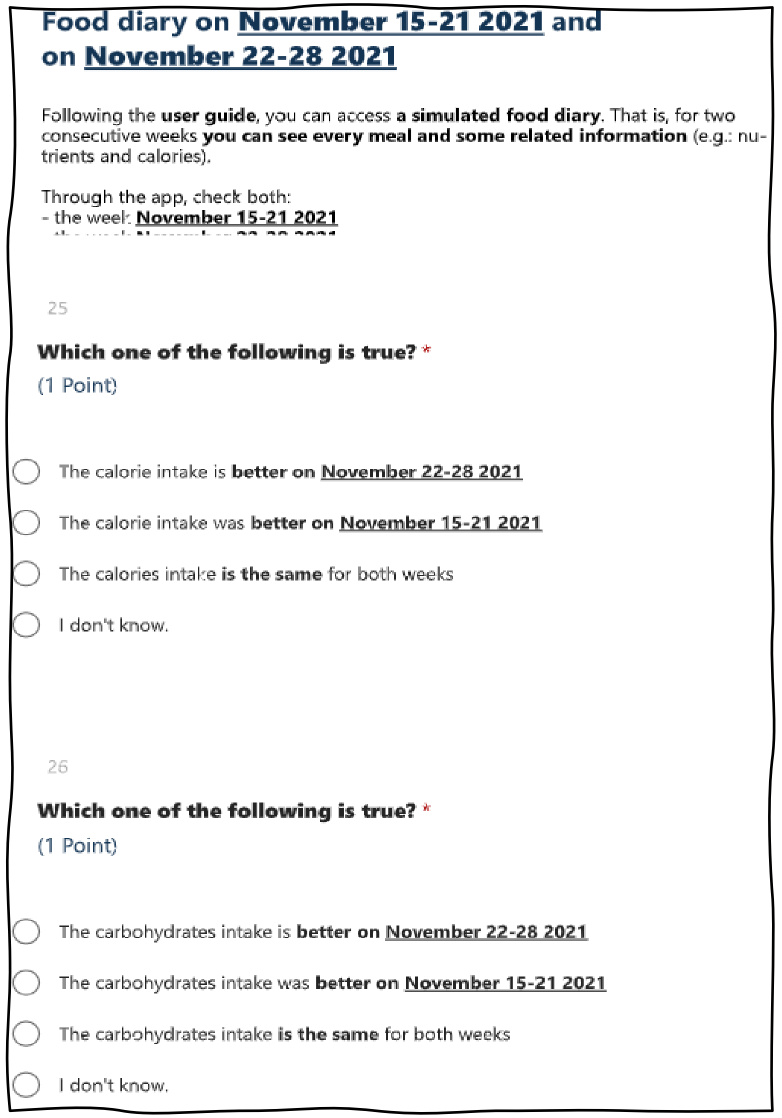
(1 Point)

- The carbohydrates intake is **too high.**
- The carbohydrates intake is **balanced.**
- The carbohydrates intake is **too low.**
- I don't know.

Measuring informativeness

- "How successfully a ~~person~~ tool is able to convey an intended message" (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet
  - **Week analysis (2pt):** intake balance across a week
  - **Weeks comparison (2pt):** detecting progress



Food diary on November 15-21 2021 and on November 22-28 2021

Following the **user guide**, you can access **a simulated food diary**. That is, for two consecutive weeks **you can see every meal and some related information** (e.g.: nutrients and calories).

Through the app, check both:
- the week: November 15-21 2021

25

Which one of the following is true? *

(1 Point)

- The calorie intake is **better on** November 22-28 2021
- The calorie intake was **better on** November 15-21 2021
- The calories intake **is the same** for both weeks
- I don't know.

26

Which one of the following is true? *

(1 Point)

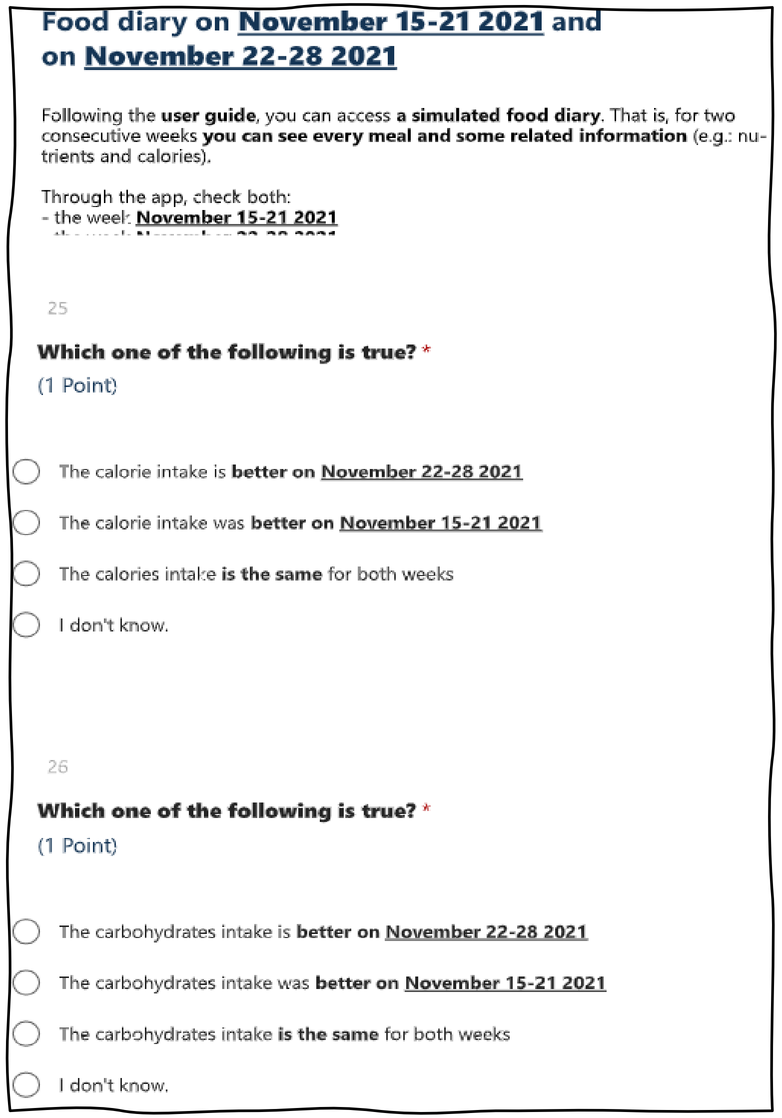
- The carbohydrates intake is **better on** November 22-28 2021
- The carbohydrates intake was **better on** November 15-21 2021
- The carbohydrates intake **is the same** for both weeks
- I don't know.

Measuring informativeness

- “How successfully a ~~person~~ tool is able to convey an intended message” (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet
  - **Week analysis (2pt):** intake balance across a week
  - **Weeks comparison (2pt):** detecting progress

- Quiz creation was guided by:



Food diary on November 15-21 2021 and on November 22-28 2021

Following the **user guide**, you can access **a simulated food diary**. That is, for two consecutive weeks **you can see every meal and some related information** (e.g.: nutrients and calories).

Through the app, check both:
- the week: November 15-21 2021

25

Which one of the following is true? *
(1 Point)

- The calorie intake is **better on** November 22-28 2021
- The calorie intake was **better on** November 15-21 2021
- The calories intake **is the same** for both weeks
- I don't know.

26

Which one of the following is true? *
(1 Point)

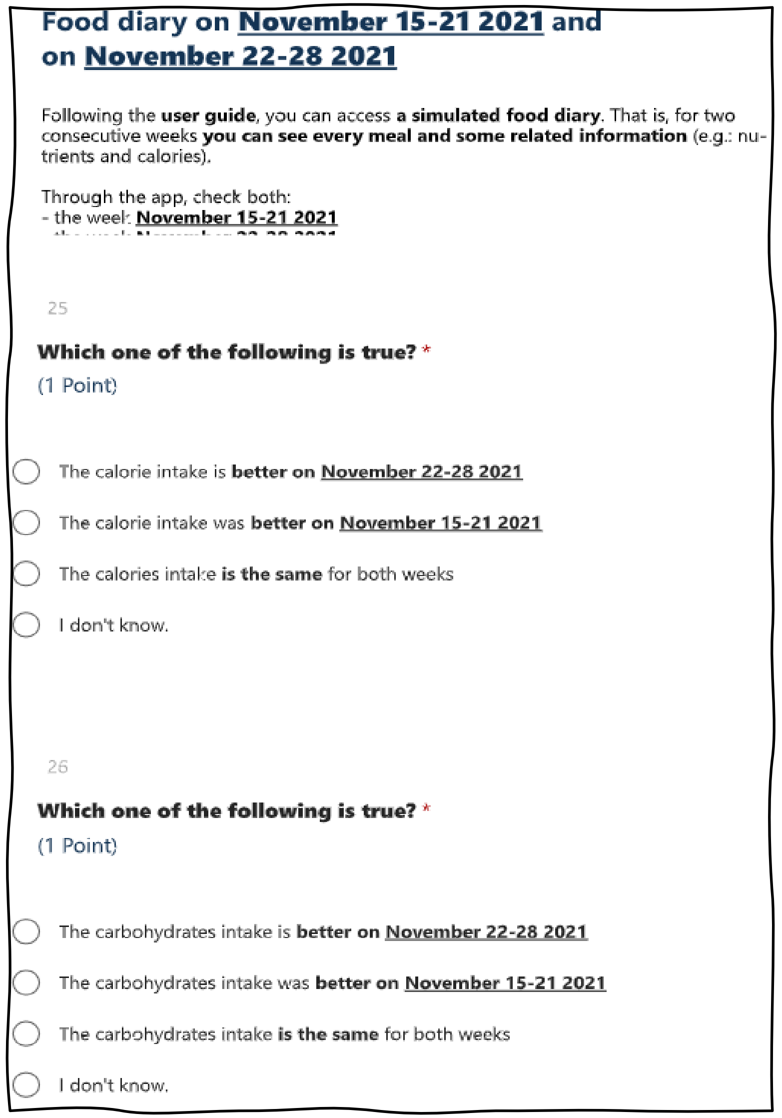
- The carbohydrates intake is **better on** November 22-28 2021
- The carbohydrates intake was **better on** November 15-21 2021
- The carbohydrates intake **is the same** for both weeks
- I don't know.

Measuring informativeness

- “How successfully a ~~person~~ tool is able to convey an intended message” (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet
  - **Week analysis (2pt):** intake balance across a week
  - **Weeks comparison (2pt):** detecting progress

- Quiz creation was guided by:
  - Expert recommendations (**Vasiloglou et al., 2020**)



Food diary on November 15-21 2021 and on November 22-28 2021

Following the **user guide**, you can access **a simulated food diary**. That is, for two consecutive weeks **you can see every meal and some related information** (e.g.: nutrients and calories).

Through the app, check both:
- the week: **November 15-21 2021**

25

Which one of the following is true? *

(1 Point)

- The calorie intake is **better on November 22-28 2021**
- The calorie intake was **better on November 15-21 2021**
- The calories intake **is the same** for both weeks
- I don't know.

26

Which one of the following is true? *

(1 Point)

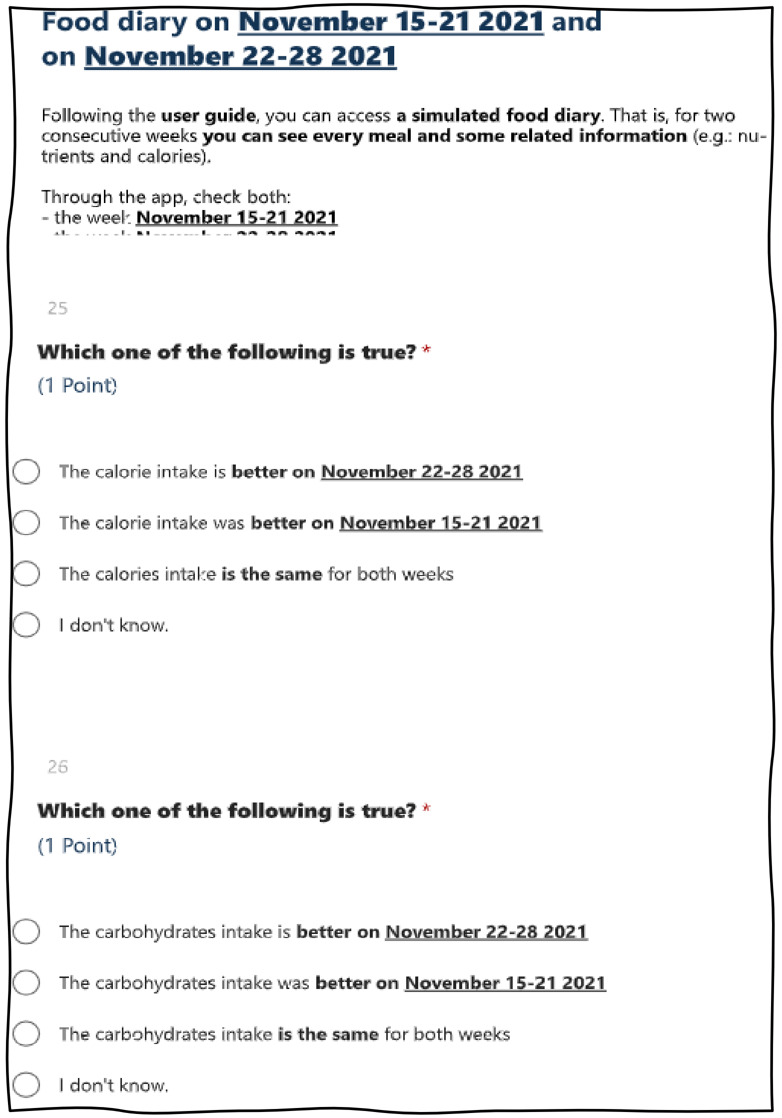
- The carbohydrates intake is **better on November 22-28 2021**
- The carbohydrates intake was **better on November 15-21 2021**
- The carbohydrates intake **is the same** for both weeks
- I don't know.

Measuring informativeness

- "How successfully a ~~person~~ tool is able to convey an intended message" (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet
  - **Week analysis (2pt):** intake balance across a week
  - **Weeks comparison (2pt):** detecting progress

- Quiz creation was guided by:
  - Expert recommendations (**Vasiloglou et al., 2020**)
  - Self-regulation principle (**Zahri et al., 2016**)



Food diary on November 15-21 2021 and on November 22-28 2021

Following the **user guide**, you can access a **simulated food diary**. That is, for two consecutive weeks **you can see every meal and some related information** (e.g.: nutrients and calories).

Through the app, check both:
- the week: **November 15-21 2021**

25

Which one of the following is true? *
(1 Point)

- The calorie intake is **better on November 22-28 2021**
- The calorie intake was **better on November 15-21 2021**
- The calories intake **is the same** for both weeks
- I don't know.

26

Which one of the following is true? *
(1 Point)

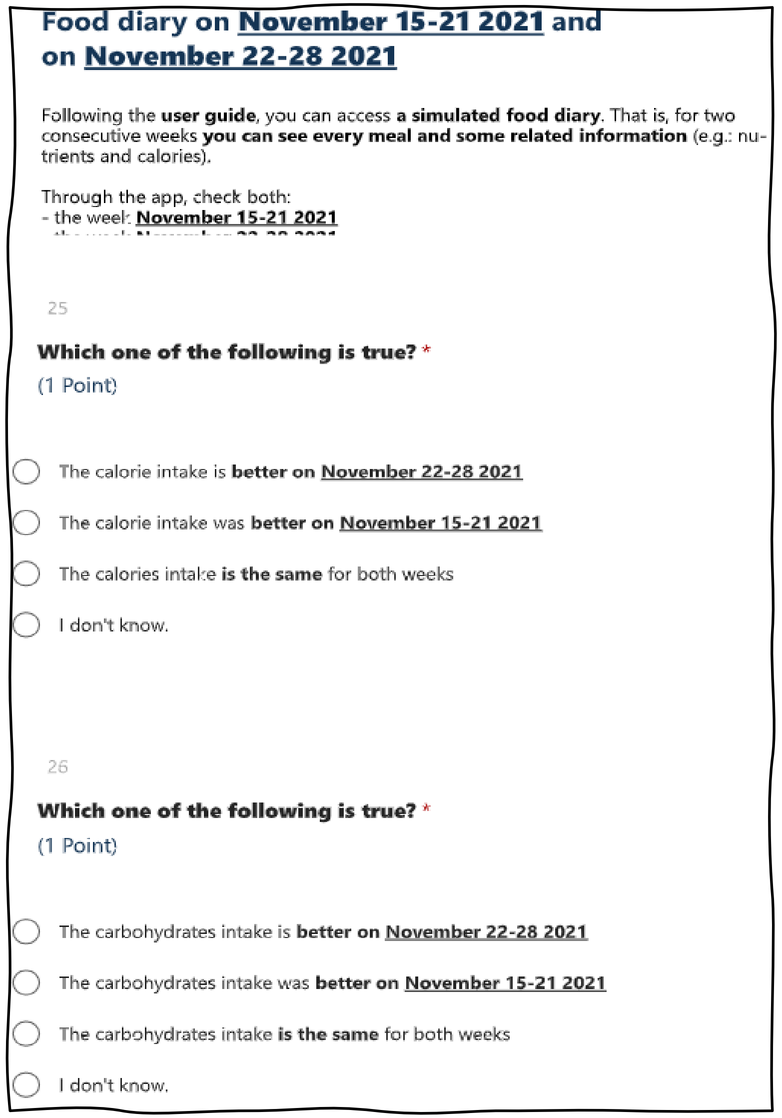
- The carbohydrates intake is **better on November 22-28 2021**
- The carbohydrates intake was **better on November 15-21 2021**
- The carbohydrates intake **is the same** for both weeks
- I don't know.

Measuring informativeness

- “How successfully a ~~person~~ tool is able to convey an intended message” (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet
  - **Week analysis (2pt):** intake balance across a week
  - **Weeks comparison (2pt):** detecting progress

- Quiz creation was guided by:
  - Expert recommendations (**Vasiloglou et al., 2020**)
  - Self-regulation principle (**Zahri et al., 2016**)
  - Existing diet apps analysis.



Food diary on November 15-21 2021 and on November 22-28 2021

Following the **user guide**, you can access a **simulated food diary**. That is, for two consecutive weeks **you can see every meal and some related information** (e.g.: nutrients and calories).

Through the app, check both:
- the week: November 15-21 2021

25

Which one of the following is true? *
(1 Point)

- The calorie intake is **better on** November 22-28 2021
- The calorie intake was **better on** November 15-21 2021
- The calories intake **is the same** for both weeks
- I don't know.

26

Which one of the following is true? *
(1 Point)

- The carbohydrates intake is **better on** November 22-28 2021
- The carbohydrates intake was **better on** November 15-21 2021
- The carbohydrates intake **is the same** for both weeks
- I don't know.

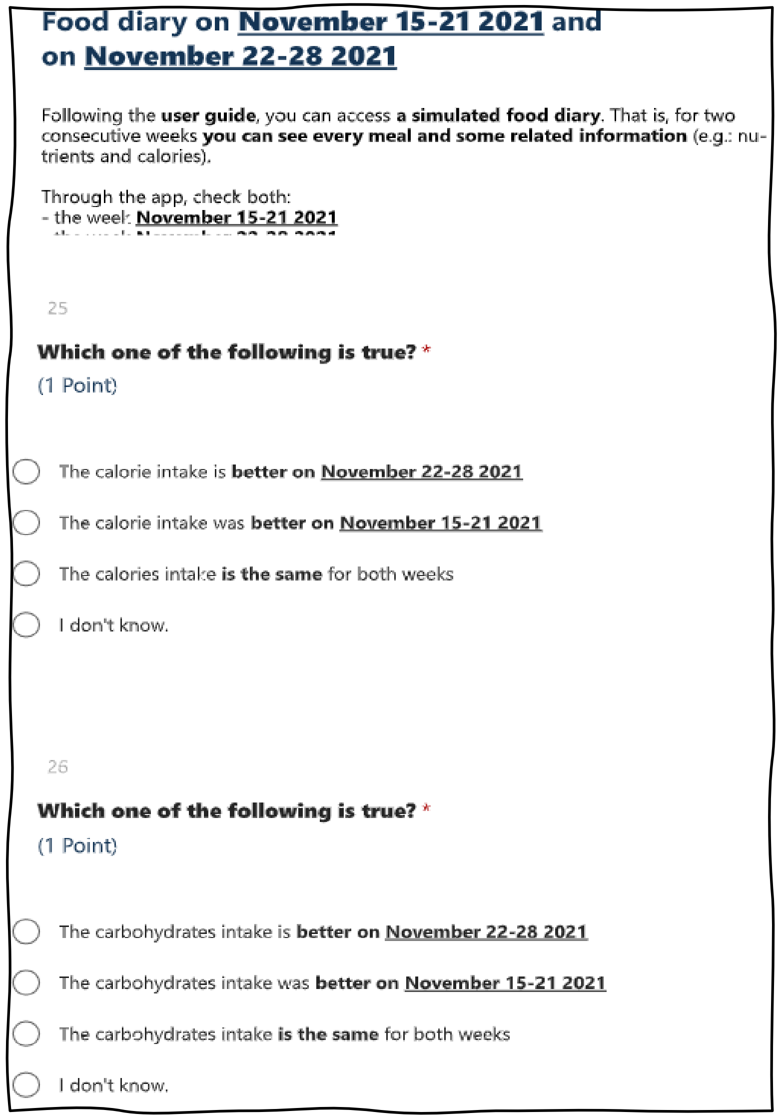
Measuring informativeness

- "How successfully a ~~person~~ tool is able to convey an intended message" (**Webster and Morris, 2019**):

- Workers used their assigned tool to take **a 10-point quiz**
  - **Day analysis (2pt):** intake balance on a single day
  - **Food analysis (4pt):** food impact on diet
  - **Week analysis (2pt):** intake balance across a week
  - **Weeks comparison (2pt):** detecting progress

- Quiz creation was guided by:
  - Expert recommendations (**Vasiloglou et al., 2020**)
  - Self-regulation principle (**Zahri et al., 2016**)
  - Existing diet apps analysis.

- H1: Chatbot workers **scored higher** on the quiz than MFP or FS workers.



Measuring nutrition literacy & feedback

Measuring nutrition literacy & feedback

- We adopt the **Newest Vital Sign** (NVS) (Weiss et al., 2005; Powers et al., 2010)

Measuring nutrition literacy & feedback

- We adopt the **Newest Vital Sign** (NVS) (**Weiss et al., 2005; Powers et al., 2010**)
  - **6 questions** about nutritional information, extracted from an ice cream label

Nutrition Facts

Serving Size	½ cup
Servings per container	4
Amount per serving	
Calories 250	Fat Cal 120
	%DV
Total Fat 13g	20%
Sat Fat 9g	40%
Cholesterol 28mg	12%
Sodium 55mg	2%
Total Carbohydrate 30g	12%
Dietary Fiber 2g	
Sugars 23g	
Protein 4g	8%

*Percentage Daily Values (DV) are based on a 2,000 calorie diet. Your daily values may be higher or lower depending on your calorie needs.

Ingredients: Cream, Skim Milk, Liquid Sugar, Water, Egg Yolks, Brown Sugar, Milkfat, Peanut Oil, Sugar, Butter, Salt, Carrageenan, Vanilla Extract.

Price: $12.72

Measuring nutrition literacy & feedback

- We adopt the **Newest Vital Sign** (NVS) (**Weiss et al., 2005; Powers et al., 2010**)
  - **6 questions** about nutritional information, extracted from an ice cream label

- H2: There was a positive correlation between **NVS score** and **quiz score** in our experiment, but not for chatbot workers.

Measuring nutrition literacy & feedback

- We adopt the **Newest Vital Sign** (NVS) (**Weiss et al., 2005; Powers et al., 2010**)
  - **6 questions** about nutritional information, extracted from an ice cream label

- H2: There was a positive correlation between **NVS score** and **quiz score** in our experiment, but not for chatbot workers.

- Finally, we ask every worker to evaluate **their assigned tool**

Please give a score to each statement, based on how much you agree with each one: *

	Disagree	Somewhat disagree	Neither agree or disagree	Somewhat agree	Agree
$tool_name helped me finding problems in the food diary.	○	○	○	○	○
$tool_name helped me answer the questions.	○	○	○	○	○
Getting the answers with **$tool_name** was quick.	○	○	○	○	○
$tool_name was easy to use and understand.	○	○	○	○	○
I think I could improve my diet using **$tool_name**.	○	○	○	○	○

Measuring nutrition literacy & feedback

- We adopt the **Newest Vital Sign** (NVS) (**Weiss et al., 2005; Powers et al., 2010**)
  - **6 questions** about nutritional information, extracted from an ice cream label

- H2: There was a positive correlation between **NVS score** and **quiz score** in our experiment, but not for chatbot workers.

- Finally, we ask every worker to evaluate **their assigned tool**
  - Additionally, we ask them about their **past experience** with dieting tools.

Please give a score to each statement, based on how much you agree with each one: *

	Disagree	Somewhat disagree	Neither agree or disagree	Somewhat agree	Agree
$tool_name helped me finding problems in the food diary.	○	○	○	○	○
$tool_name helped me answer the questions.	○	○	○	○	○
Getting the answers with **$tool_name** was quick.	○	○	○	○	○
$tool_name was easy to use and understand.	○	○	○	○	○
I think I could improve my diet using **$tool_name**.	○	○	○	○	○

Measuring nutrition literacy & feedback

- We adopt the **Newest Vital Sign** (NVS) (**Weiss et al., 2005; Powers et al., 2010**)
  - **6 questions** about nutritional information, extracted from an ice cream label
- H2: There was a positive correlation between **NVS score** and **quiz score** in our experiment, but not for chatbot workers.
- Finally, we ask every worker to evaluate **their assigned tool**
  - Additionally, we ask them about their **past experience** with dieting tools.
- H3: Our chatbot received **better ratings** than MFP and FS

Please give a score to each statement, based on how much you agree with each one: *

	Disagree	Somewhat disagree	Neither agree or disagree	Somewhat agree	Agree
$tool_name helped me finding problems in the food diary.	○	○	○	○	○
$tool_name helped me answer the questions.	○	○	○	○	○
Getting the answers with **$tool_name** was quick.	○	○	○	○	○
$tool_name was easy to use and understand.	○	○	○	○	○
I think I could improve my diet using **$tool_name**.	○	○	○	○	○

Preliminary checks

Preliminary checks

- Nutrition literacy uniformity

	Workers per class		
NVS class	CB	FS	MFP
LOW (0-1pt)	1	0	9
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with **low NVS scores** are in MFP sample

	Workers per class		
NVS class	CB	FS	MFP
LOW (0-1pt)	1	0	9
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with **low NVS scores** are in MFP sample
  - Solution: **removing** every worker with low NVS score.

	Workers per class		
NVS class	CB	FS	MFP
~~LOW (0-1pt)~~	~~1~~	~~0~~	~~9~~
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with low NVS scores are in MFP sample
  - Solution: removing every worker with low NVS score.
    - Pro: keeps the comparison fair

	Workers per class		
NVS class	CB	FS	MFP
~~LOW (0-1pt)~~	~~1~~	~~0~~	~~9~~
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with **low NVS scores** are in MFP sample
  - Solution: **removing** every worker with low NVS score.
    - Pro: keeps the comparison fair
    - Cons: potentially invalidates H2

	Workers per class		
NVS class	CB	FS	MFP
~~LOW (0-1pt)~~	~~1~~	~~0~~	~~9~~
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with **low NVS scores** are in MFP sample
  - Solution: **removing** every worker with low NVS score.
    - Pro: keeps the comparison fair
    - Cons: potentially invalidates H2
- Comparison of past experience with diet apps

	Workers per class		
NVS class	CB	FS	MFP
~~LOW (0-1pt)~~	~~1~~	~~0~~	~~9~~
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with **low NVS scores** are in MFP sample
  - Solution: **removing** every worker with low NVS score.
    - Pro: keeps the comparison fair
    - Cons: potentially invalidates H2

- Comparison of past experience with diet apps
  - No significant difference either:

	Workers per class		
NVS class	CB	FS	MFP
~~LOW (0-1pt)~~	~~1~~	~~0~~	~~9~~
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with **low NVS scores** are in MFP sample
  - Solution: **removing** every worker with low NVS score.
    - Pro: keeps the comparison fair
    - Cons: potentially invalidates H2
- Comparison of past experience with diet apps
  - No significant difference either:
    - Globally **(p=0.47)**

	Workers per class		
NVS class	CB	FS	MFP
~~LOW (0-1pt)~~	~~1~~	~~0~~	~~9~~
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Preliminary checks

- Nutrition literacy uniformity
  - Problem: most workers with **low NVS scores** are in MFP sample
  - Solution: **removing** every worker with low NVS score.
    - Pro: keeps the comparison fair
    - Cons: potentially invalidates H2

- Comparison of past experience with diet apps
  - No significant difference either:
    - Globally **(p=0.47)**
    - For workers who used diet tools in the past only **(p=0.27)**

	Workers per class		
NVS class	CB	FS	MFP
~~LOW (0-1pt)~~	~~1~~	~~0~~	~~9~~
MID (2-3pt)	5	3	5
HIGH (4-6pt)	21	26	17

Table 2: Distribution of nutrition literacy for our population.

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, CB beat both apps

Topic	Average score		
	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

Topic	Score differences		
	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, **CB** beat both apps
  - Significantly better than **FS**

Topic	Average score CB	Average score FS	Average score MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

Topic	Score differences CB-FS	Score differences CB-MFP	Score differences MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, CB beat both apps
  - Significantly better than FS
  - Regardless of the tool, scores are low

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, **CB** beat both apps
  - Significantly better than **FS**
  - Regardless of the tool, **scores are low**

- Per-task: **CB** beat both apps **in most cases**

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, **CB** beat both apps
  - Significantly better than **FS**
  - Regardless of the tool, **scores are low**

- Per-task: **CB** beat both apps **in most cases**
  - Significantly better than **MFP** in:
    - Food analysis

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, **CB** beat both apps
  - Significantly better than **FS**
  - Regardless of the tool, **scores are low**

- Per-task: **CB** beat both apps **in most cases**
  - Significantly better than **MFP** in:
    - Food analysis
  - Significantly better than **FS** in:

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, **CB** beat both apps
  - Significantly better than **FS**
  - Regardless of the tool, **scores are low**

- Per-task: **CB** beat both apps **in most cases**
  - Significantly better than **MFP** in:
    - Food analysis
  - Significantly better than **FS** in:
    - Week analysis
    - Weeks comparison

Topic	Average score		
	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

Topic	Score differences		
	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, **CB** beat both apps
  - Significantly better than **FS**
  - Regardless of the tool, **scores are low**

- Per-task: **CB** beat both apps **in most cases**
  - Significantly better than **MFP** in:
    - Food analysis
  - Significantly better than **FS** in:
    - Week analysis
    - Weeks comparison
  - CB was beaten only by **MFP** in day analysis

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, **CB** beat both apps
  - Significantly better than **FS**
  - Regardless of the tool, **scores are low**

- Per-task: **CB** beat both apps **in most cases**
  - Significantly better than **MFP** in:
    - Food analysis
  - Significantly better than **FS** in:
    - Week analysis
    - Weeks comparison
  - CB was beaten only by **MFP** in day analysis

- Is there a link between the used tool and scores in specific tasks?

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Overall, CB beat both apps
  - Significantly better than FS
  - Regardless of the tool, scores are low

- Per-task: CB beat both apps in most cases
  - Significantly better than MFP in:
    - Food analysis
  - Significantly better than FS in:
    - Week analysis
    - Weeks comparison
  - CB was beaten only by MFP in day analysis

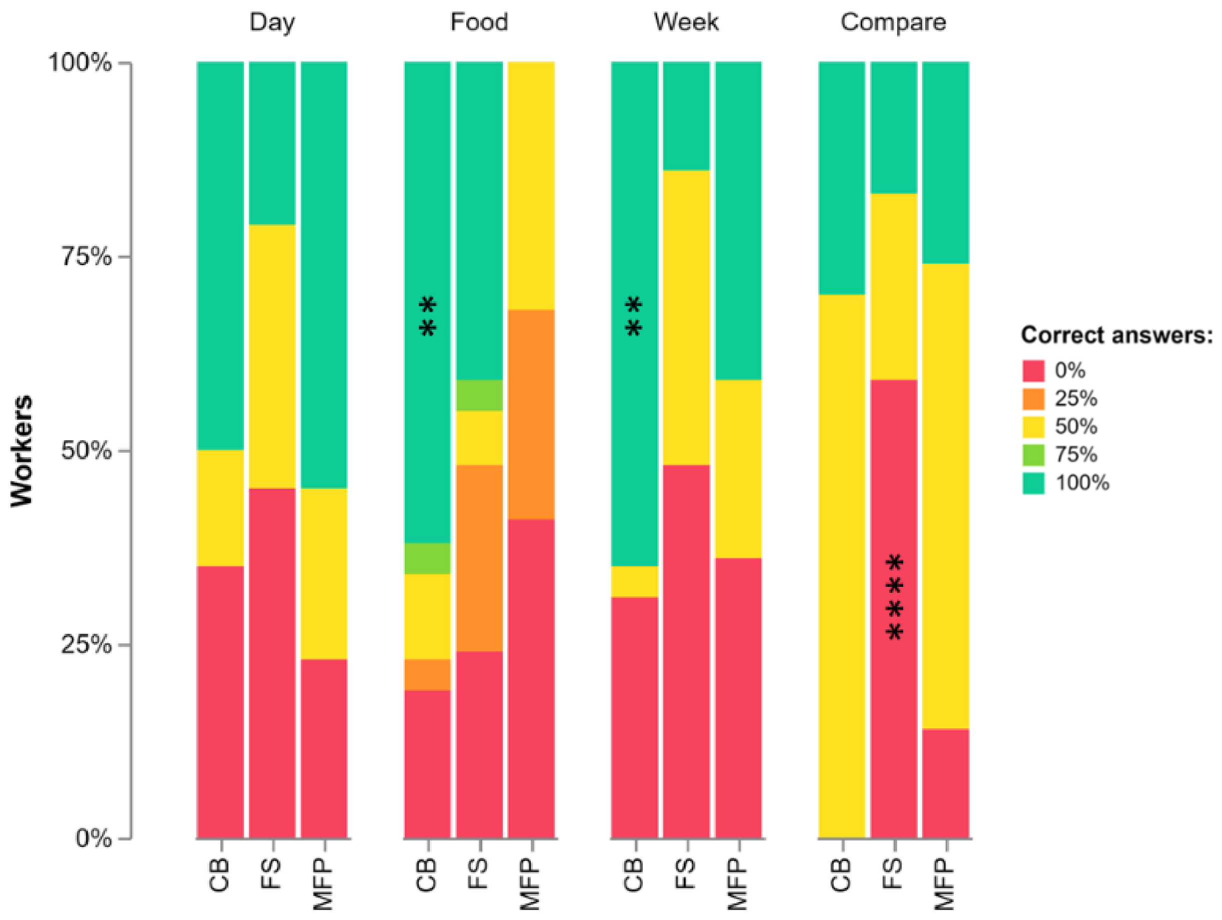
- Is there a link between the used tool and scores in specific tasks?
  - We inspect score distribution by tool and tasks

	Average score		
Topic	CB	FS	MFP
Overall (10pt)	**6.65**	4.13	5.22
Day analysis (2pt)	1.15	0.76	**1.32**
Food analysis (4pt)	**2.85**	2.14	0.91
Week analysis (2pt)	**1.35**	0.66	1.05
Weeks comparison (2pt)	**1.31**	0.59	1.14

	Score differences		
Topic	CB-FS	CB-MFP	MFP-FS
Overall (10pt)	+2.52**	+1.43	+1.09
Day analysis (2pt)	+0.40	-0.16	+0.56
Food analysis (4pt)	+0.71	+1.94***	-1.23*
Week analysis (2pt)	+0.70**	+0.30	+0.39
Weeks comparison (2pt)	+0.72**	+0.17	+0.55**

CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

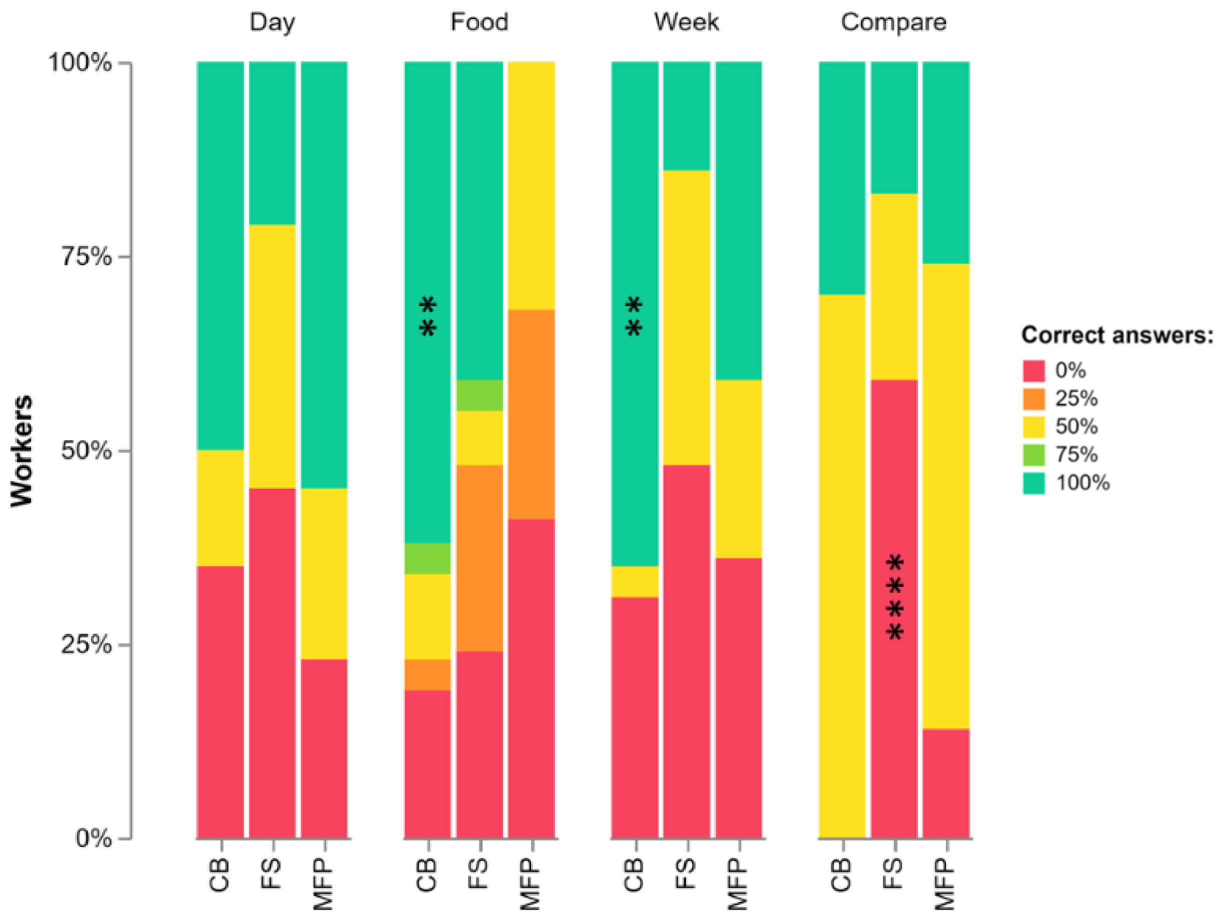
Quiz scores



CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

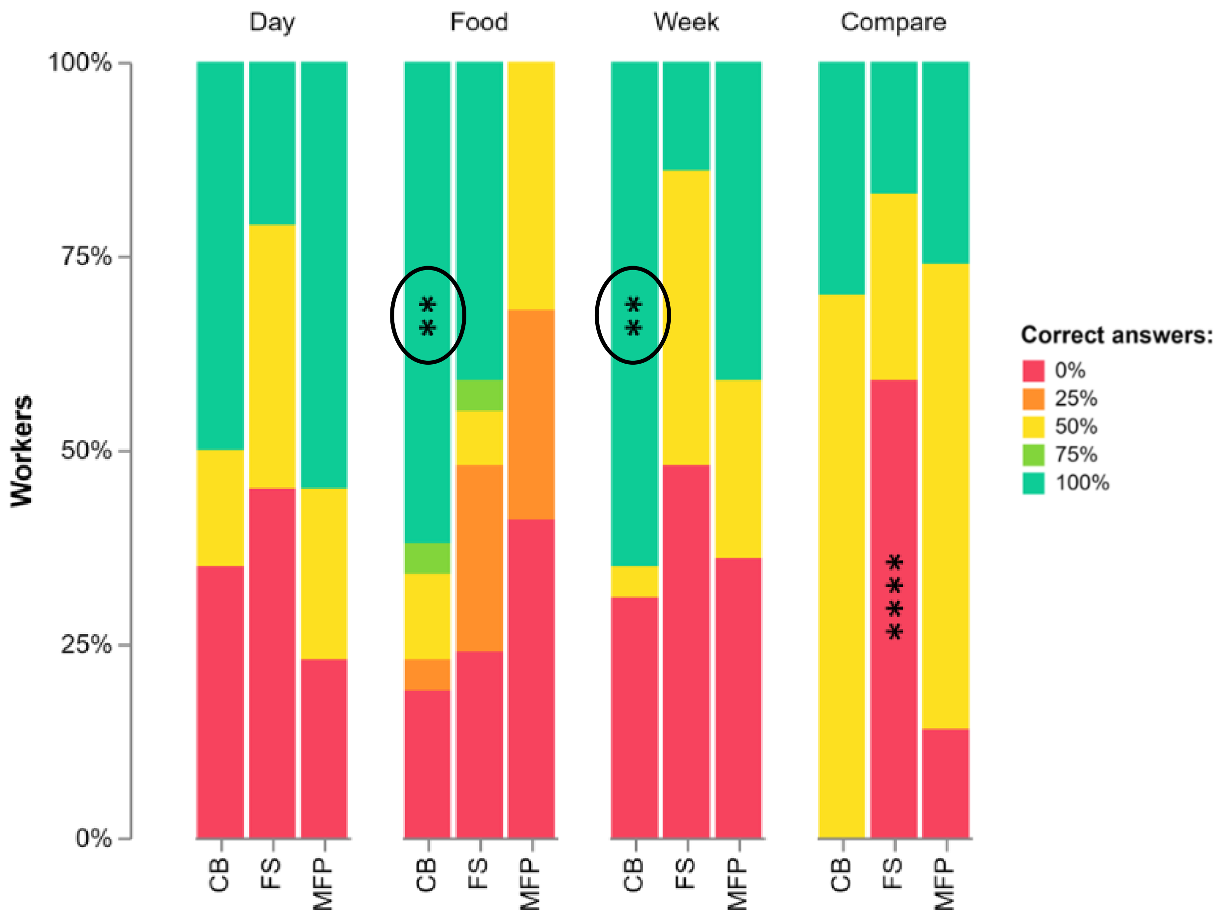
- Results for CB:



CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

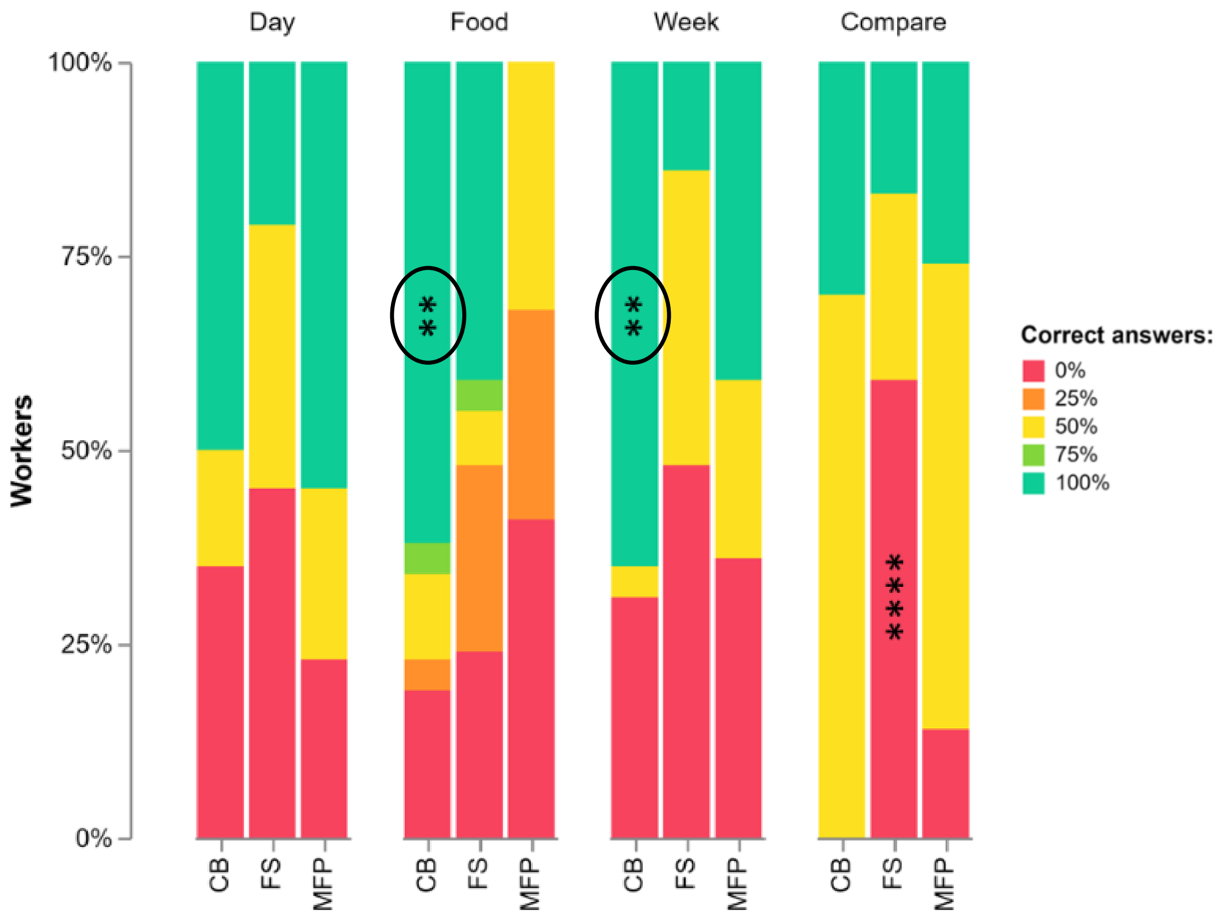
- Results for CB:
  - Positive association with:
    - Scoring 100% in food analysis
    - Scoring 100% in week analysis



CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Results for CB:
  - Positive association with:
    - Scoring 100% in food analysis
    - Scoring 100% in week analysis
  - Negative association with:
    - Scoring 0% in weeks comparisons

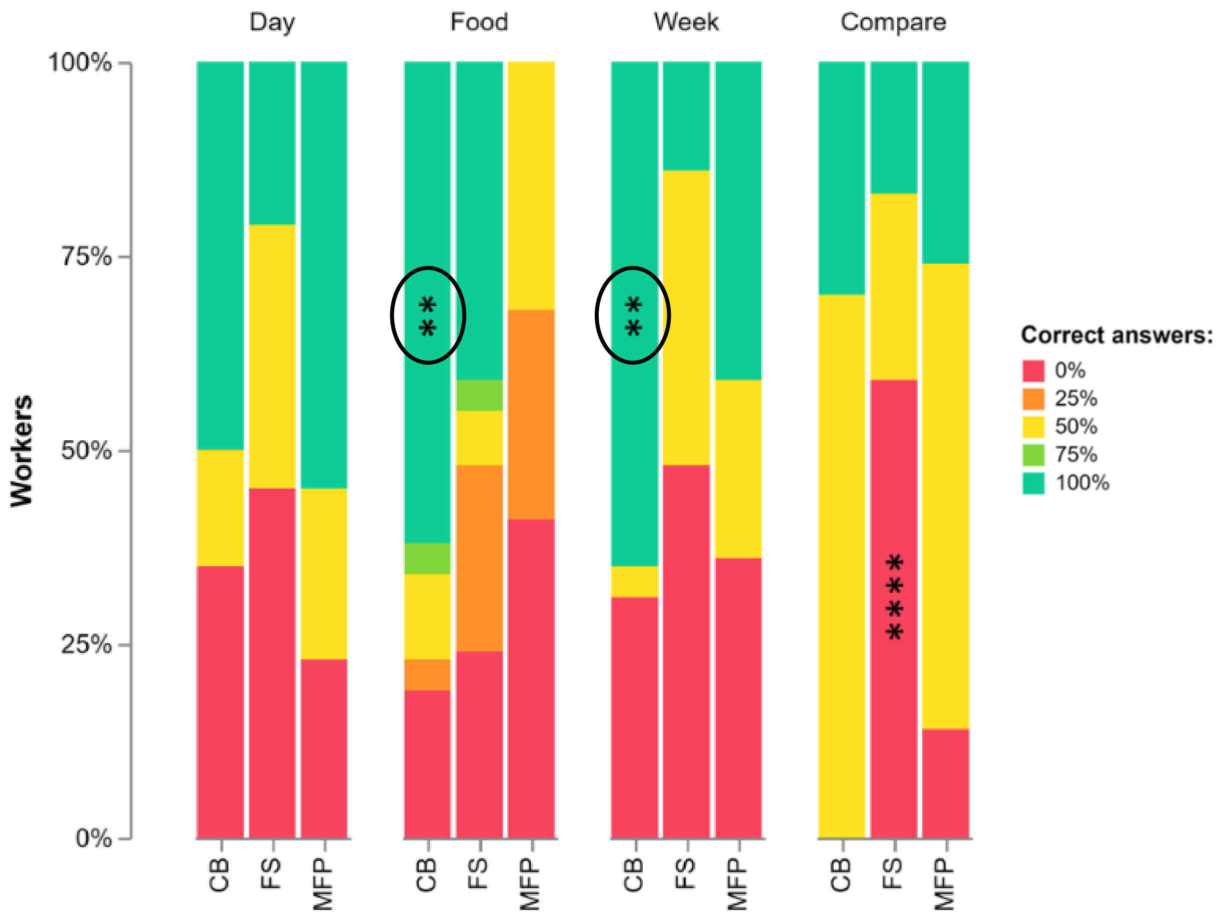


CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Results for CB:
  - Positive association with:
    - Scoring 100% in food analysis
    - Scoring 100% in week analysis
  - Negative association with:
    - Scoring 0% in weeks comparisons

- Results for FS:

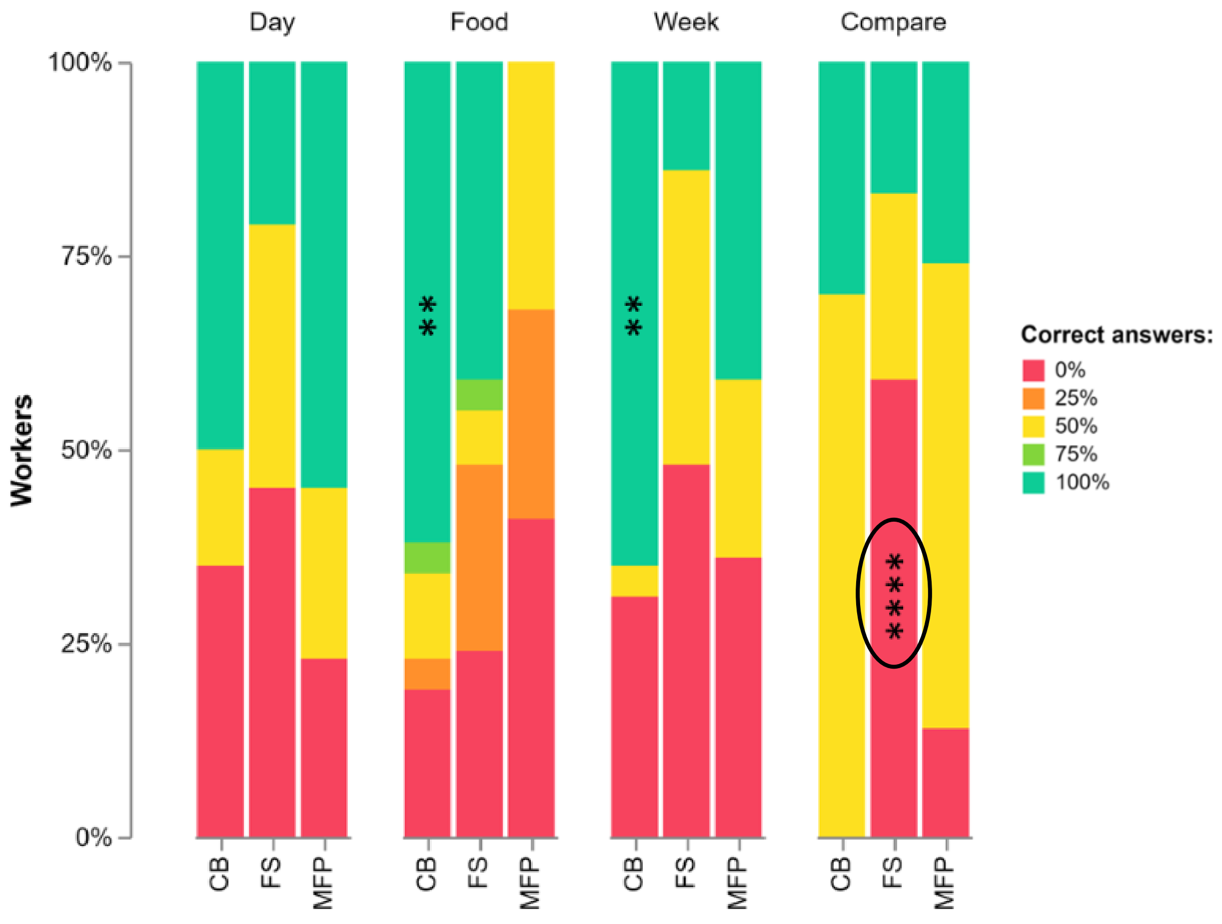


CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Results for CB:
  - Positive association with:
    - Scoring 100% in food analysis
    - Scoring 100% in week analysis
  - Negative association with:
    - Scoring 0% in weeks comparisons

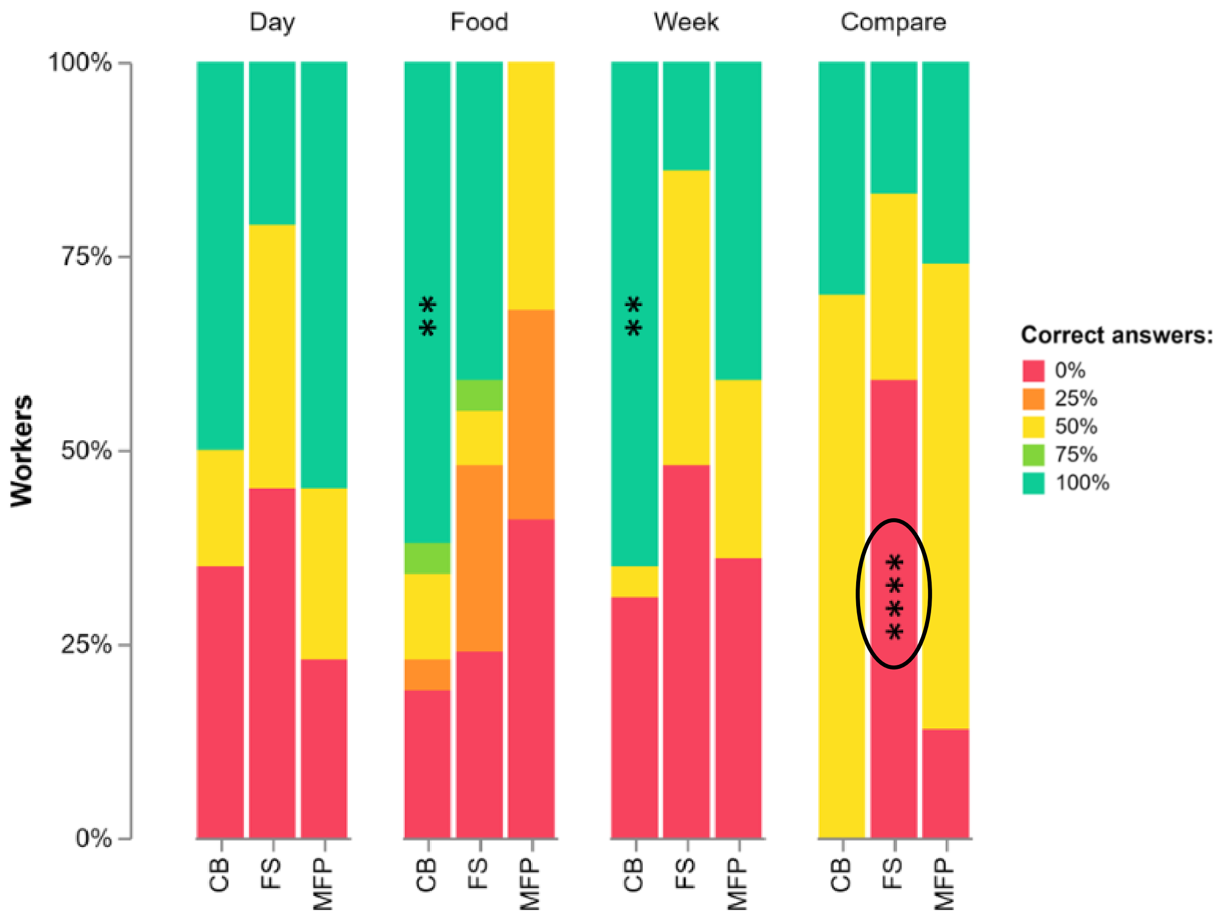
- Results for FS:
  - Positive association with:
    - Scoring 0% in weeks comparison



CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Results for CB:
  - Positive association with:
    - Scoring 100% in food analysis
    - Scoring 100% in week analysis
  - Negative association with:
    - Scoring 0% in weeks comparisons
- Results for FS:
  - Positive association with:
    - Scoring 0% in weeks comparison
- Results for MFP:



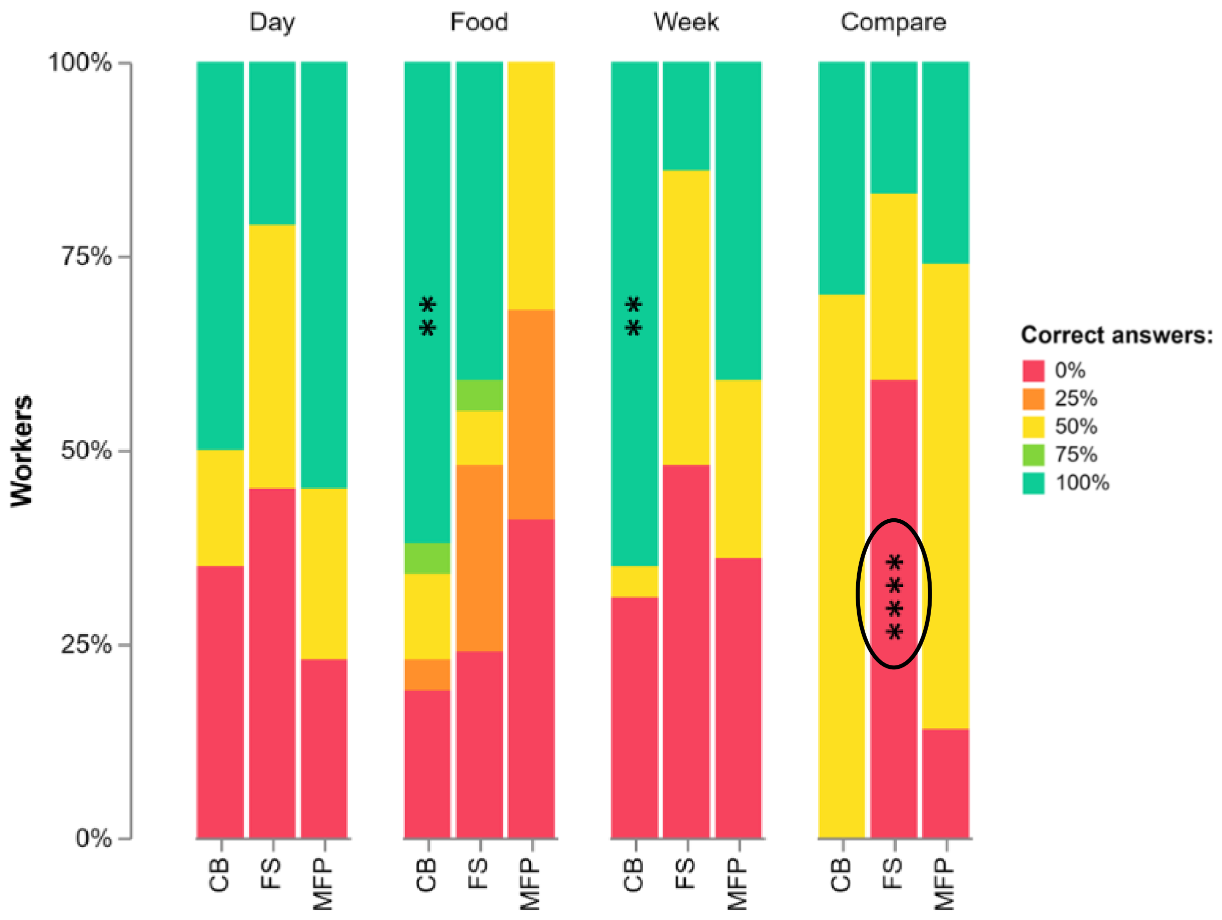
CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Quiz scores

- Results for CB:
  - Positive association with:
    - Scoring 100% in food analysis
    - Scoring 100% in week analysis
  - Negative association with:
    - Scoring 0% in weeks comparisons

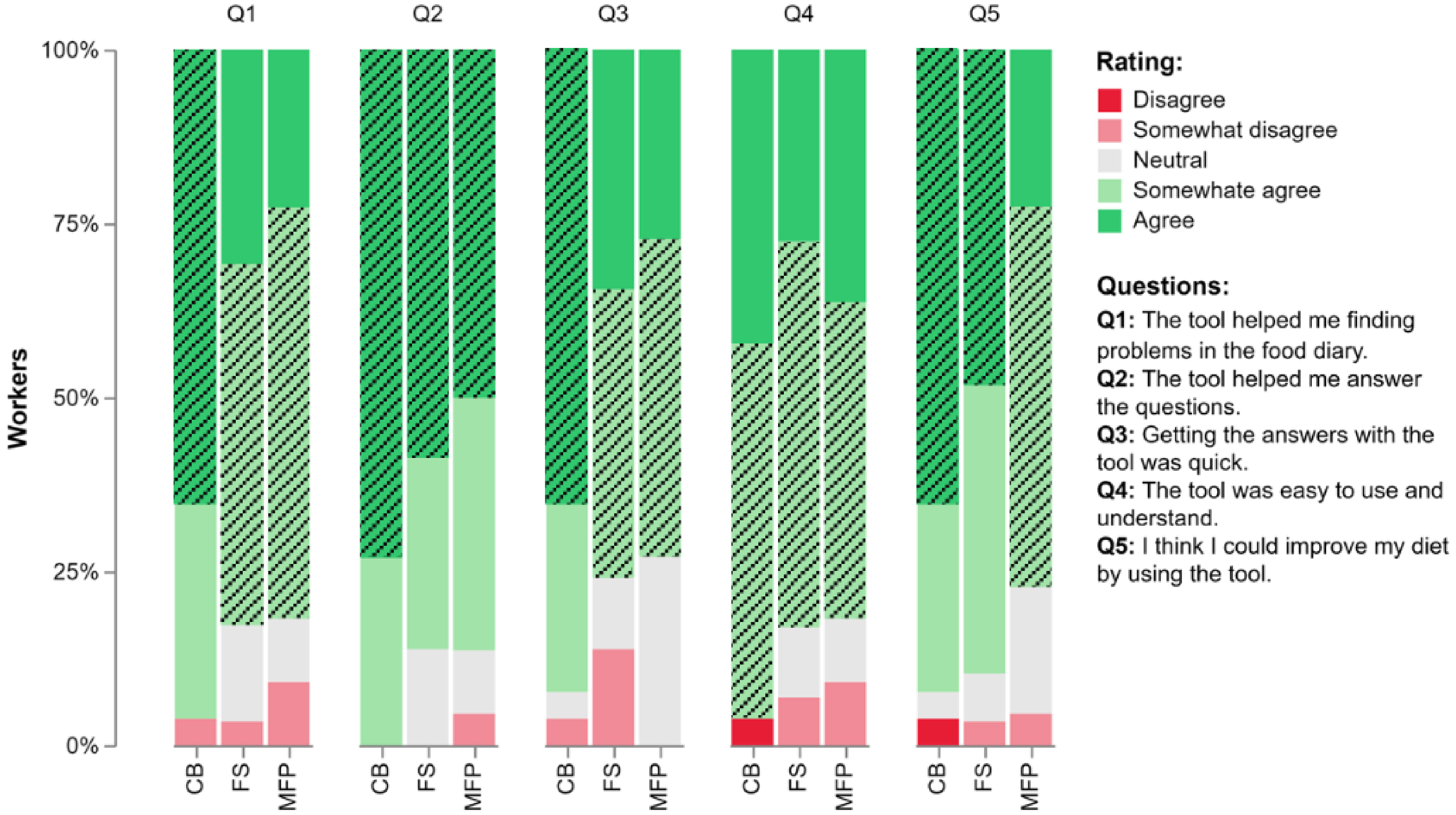
- Results for FS:
  - Positive association with:
    - Scoring 0% in weeks comparison

- Results for MFP:
  - Negative association with:
    - Scoring 100% in food analysis



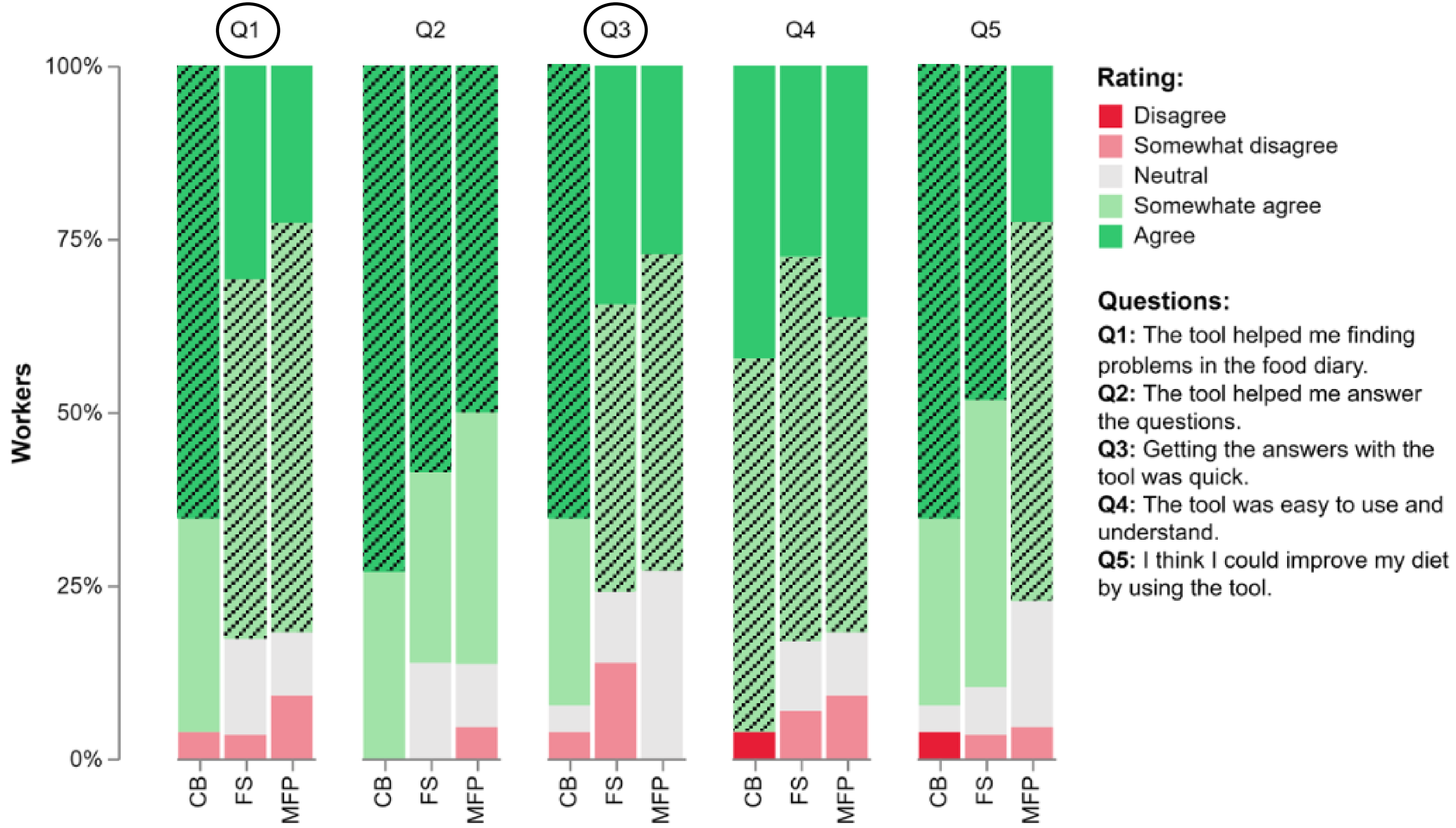
CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Perception of the tool



CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Perception of the tool



CB = Chatbot; FS = FatSecret; MFP = MyFitnessPal

Conclusions

Conclusions

- We evaluated the effect of text and charts combination on informativeness in the conversational scenario, for diet coaching.

Conclusions

- We evaluated the effect of text and charts combination on informativeness in the conversational scenario, for diet coaching.

- Our results on informativeness:

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

- Our results on nutrition literacy:

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

- Our results on nutrition literacy:
  - **We saw a relationship** between nutrition literacy and quiz scores

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

- Our results on nutrition literacy:
  - **We saw a relationship** between nutrition literacy and quiz scores
  - Our sample was **too unbalanced to test** whether our chatbot could alleviate this

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

- Our results on nutrition literacy:
  - **We saw a relationship** between nutrition literacy and quiz scores
  - Our sample was **too unbalanced to test** whether our chatbot could alleviate this
  - H2: neither confirmed nor rejected

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

- Our results on nutrition literacy:
  - **We saw a relationship** between nutrition literacy and quiz scores
  - Our sample was **too unbalanced to test** whether our chatbot could alleviate this
  - H2: neither confirmed nor rejected

- Our chatbot consistently received **better ratings** than static dieting apps

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

- Our results on nutrition literacy:
  - **We saw a relationship** between nutrition literacy and quiz scores
  - Our sample was **too unbalanced to test** whether our chatbot could alleviate this
  - H2: neither confirmed nor rejected

- Our chatbot consistently received **better ratings** than static dieting apps
- Our chatbot was considered **faster and more useful**

Conclusions

- We evaluated the effect of text and charts combination **on informativeness** in the conversational scenario, for diet coaching.

- Our results on informativeness:
  - Our chatbot consistently **outperformed static dieting apps**
  - Our chatbot was **associated with high scores**; the opposite was found for apps
  - H1: confirmed

- Our results on nutrition literacy:
  - **We saw a relationship** between nutrition literacy and quiz scores
  - Our sample was **too unbalanced to test** whether our chatbot could alleviate this
  - H2: neither confirmed nor rejected

- Our chatbot consistently received **better ratings** than static dieting apps
- Our chatbot was considered **faster and more useful**
- H3: confirmed

Thank you!

Questions?

1495 UNIVERSITY OF ABERDEEN

Horizon 2020
European Union Funding
for Research & Innovation

PHIL HUMANS