

Praxisbeispiele Prüfen mit ChatGPT

Erkenntnisse aus dem PgB8 Projekt

*Untersuchung von Chancen und Limitationen
des Einsatzes von ChatGPT in verschiedensten
Prüfungsformen in der technischen Ausbildung*

April 23- Mai 24

Schlatter Monika
30.11.23



Fragestellungen des Projektes

Wie verwenden Studierende ChatGPT in verschiedenen Prüfungssettings?

Wie schätzen sie dabei die Nützlichkeit von ChatGPT ein?

Welche Limitationen sehen sie?

Was ziehen sie daraus für Schlüsse?

Was beobachten Dozierende bei der Verwendung von ChatGPT in ihren Prüfungssettings?

Welche Schlüsse können wir daraus für die Gestaltung kommender Prüfungen ziehen?

ChatGPT als Hilfsmittel in 4 verschiedenen Settings

Semesterarbeit

E-Prüfung

mündliche Prüfung

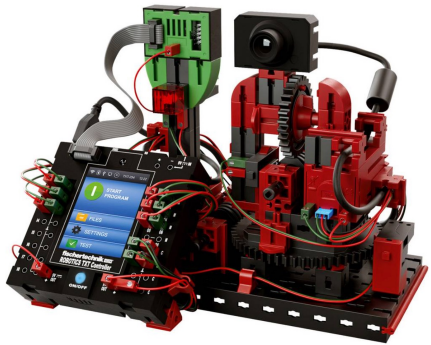
Gruppenarbeit

Smart Factory

Datenbankdesign & Implementation

Web Datenbeschaffung

Grundverständnis Data Science, Datenvisualisierung und Explorative Datenanalyse



```

Query Query History
1 CREATE TABLE author_info AS
2 SELECT *
3 FROM author_details;
    
```

```

import re
hand = open('mbox-short.txt')
for line in hand:
    line = line.rstrip()
    if re.search('From:', line):
        print(line)
    
```

- Pakete aktivieren (Pakete `tidyverse`, `tidymodels`, mittels Funktionen `install.packages()`, `library()`)
- Daten einlesen (Pfad der Datenablage definieren und F)
- Daten anschauen und verstehen (Funktion `glimpse()`)
- Datenstruktur und -typen untersuchen (Funktion `str()`)
- Neue Attribute erzeugen für (a) Anzahl Familienmitglieder (`mutate()`, `case_when()` verwenden)

Konzepterstellung
Modellerstellung
Programmieren (grafisch)

Modellverständnis
Programmieren

Programmieren
Verständnisfragen

Programmieren
Visualisieren
Konzepte

Wie verwenden Studierende ChatGPT?

Wie schätzen sie die Nützlichkeit von ChatGPT ein?

Welche Limitationen sehen sie?

Was ziehen sie daraus für Schlüsse?

Wieviele Studierende haben ChatGPT genutzt?

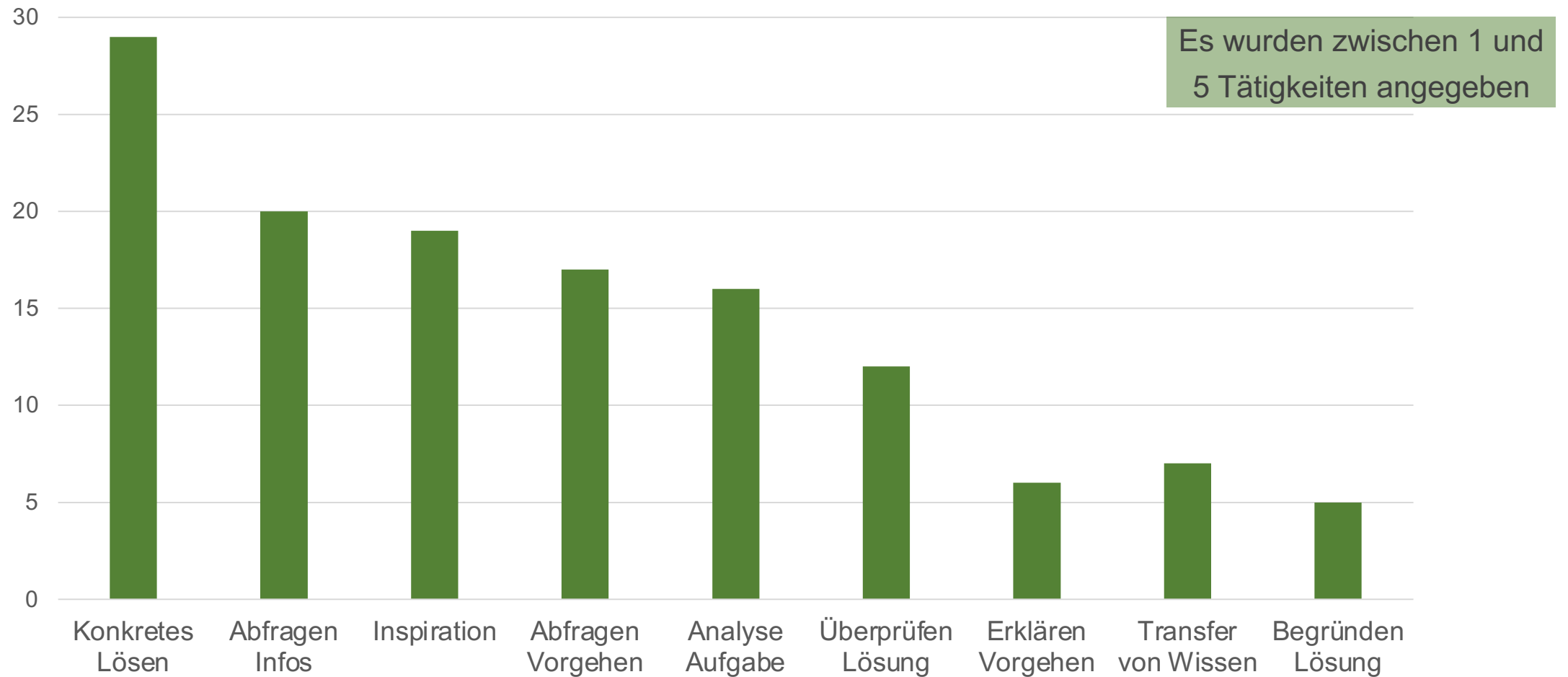
	Semesterarbeit	E- Prüfung	Mündliche Prüfung	Gruppenarbeit
# Studierende	18	15	8	37
Rücklauf Umfrage	100%	40%	nur ChatGPT Nutzende haben geantwortet	73%
Studierende, die ChatGPT genutzt haben	83% (n=15)	83% (n=5)	88% (n=7)	50% (n=14)

Für welche Tätigkeiten wurde ChatGPT verwendet?

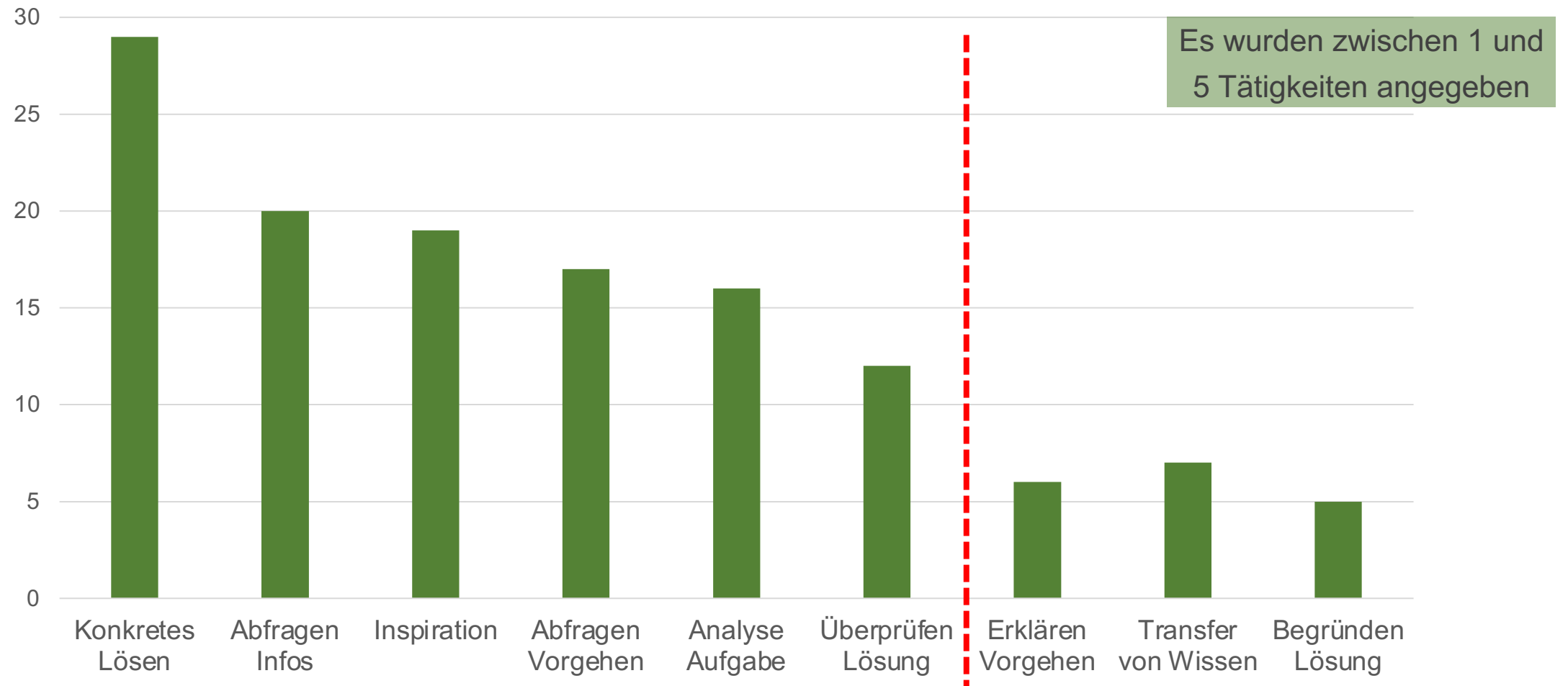
Mehrfachnennungen möglich

- Analysieren der gestellten Aufgabe (was bedeutet diese Aufgabe überhaupt?)
- Abfragen von Informationen und unklaren Begriffen (was ist/heisst...)
- Inspiration, wie die Aufgabe überhaupt gelöst werden konnte (wie könnte ich...)
- Abfragen einer konkreten Vorgehensweise zur Lösung der Aufgabe (wie mache ich...)
- Erklären der Vorgehensweise (warum so...)
- Konkretes Lösen der Aufgabe oder Teilschritte davon (gib mir eine genaue Anleitung/schreib mir Code/....)
- Transfer deines Wissens in die konkrete vorliegende Aufgabe (ich weiss X und Y. Wie wende ich dies auf diese Aufgabe an?)
- Überprüfen der Lösung/des Vorgehens (stimmt dies so?)
- Begründen der Vorgehensweise und Lösung (warum stimmt dies so?)

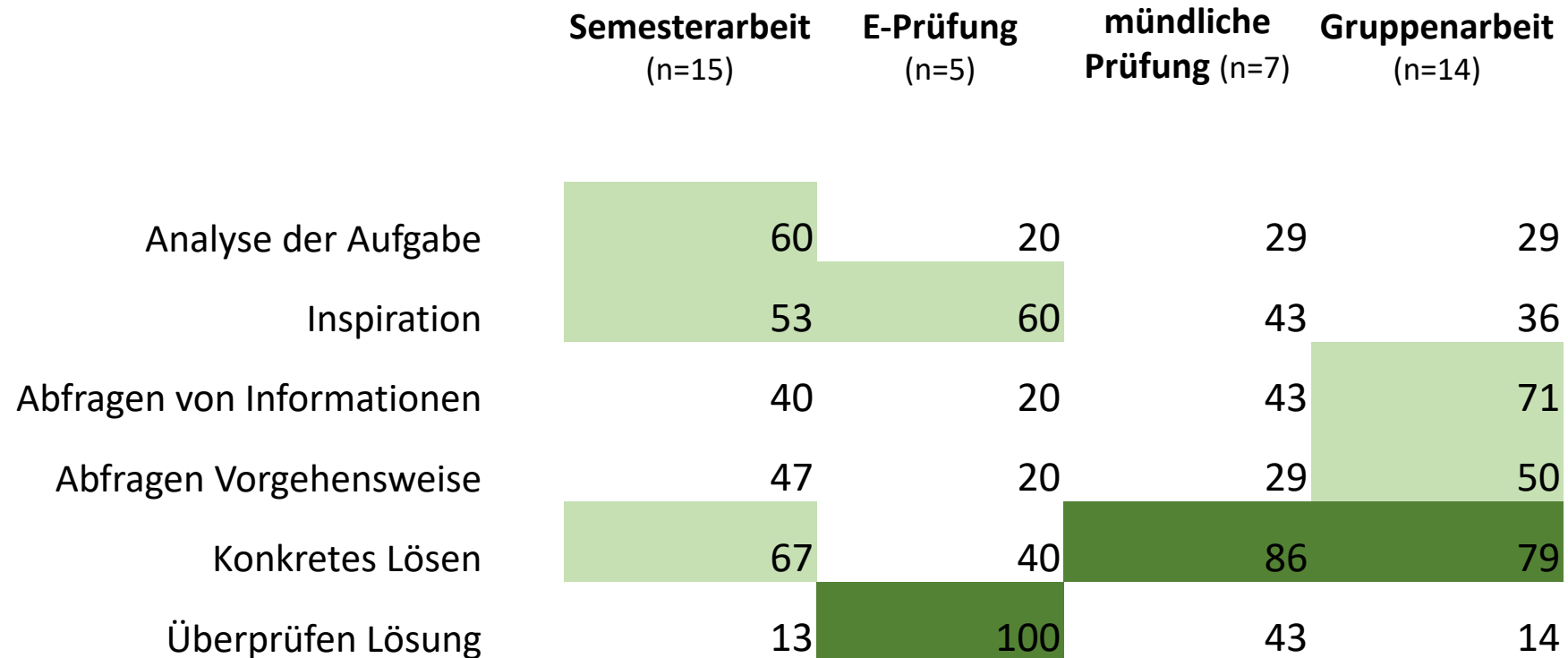
Anzahl Verwendungen alle Studierende (n=41)



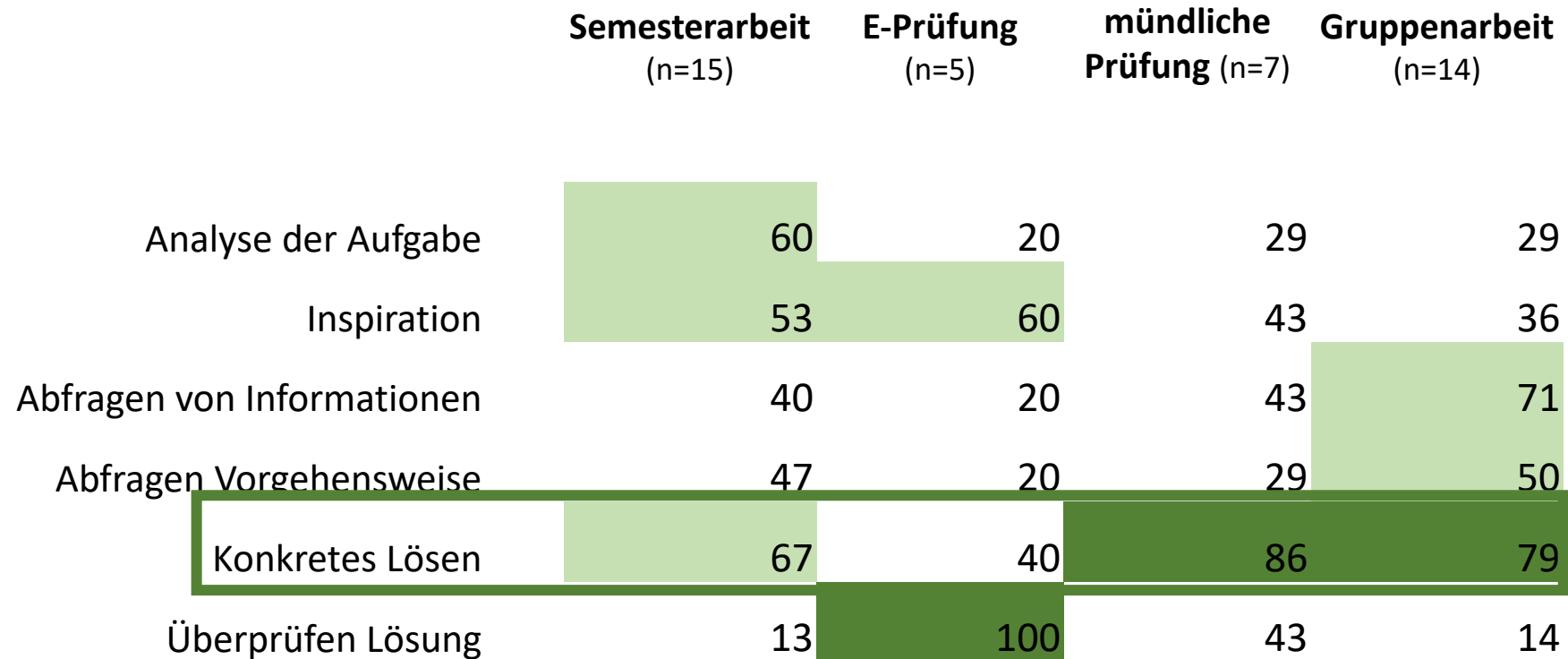
Anzahl Verwendungen alle Studierende (n=41)



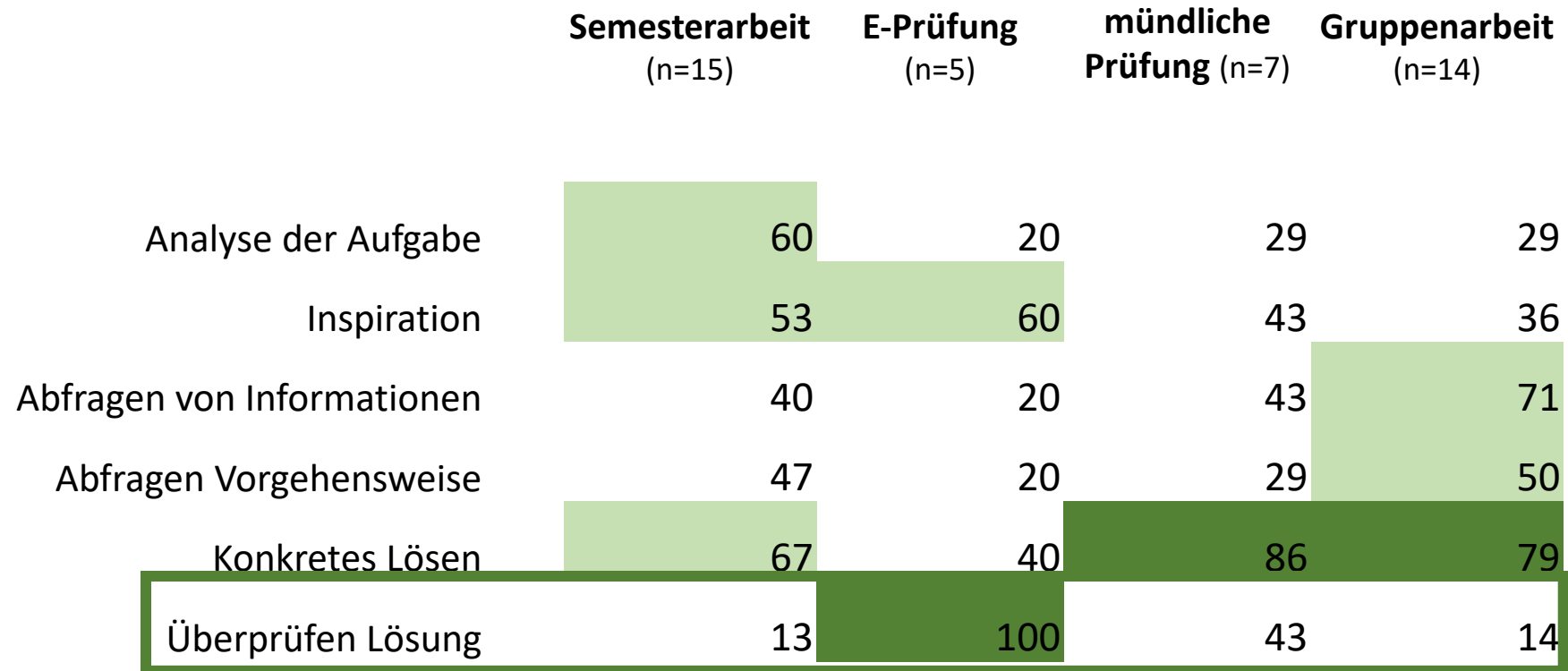
% Studierende im Modul, die jeweils ChatGPT verwendet haben



% Studierende im Modul, die jeweils ChatGPT verwendet haben



% Studierende im Modul, die jeweils ChatGPT verwendet haben



% Studierende im Modul, die jeweils ChatGPT verwendet haben

	Semesterarbeit (n=15)	E-Prüfung (n=5)	mündliche Prüfung (n=7)	Gruppenarbeit (n=14)
Analyse der Aufgabe	60	20	29	29
Inspiration	53	60	43	36
Abfragen von Informationen	40	20	43	71
Abfragen Vorgehensweise	47	20	29	50
Konkretes Lösen	67	40	86	79
Überprüfen Lösung	13	100	43	14

% Studierende im Modul, die jeweils ChatGPT verwendet haben

	Semesterarbeit (n=15)	E-Prüfung (n=5)	mündliche Prüfung (n=7)	Gruppenarbeit (n=14)
Analyse der Aufgabe	60	20	29	29
Inspiration	53	60	43	36
Abfragen von Informationen	40	20	43	71
Abfragen Vorgehensweise	47	20	29	50
Konkretes Lösen	67	40	86	79
Überprüfen Lösung	13	100	43	14

Zusammenfassung

Lösen: Hauptverwendung, sehr ausgeprägt bei Prüfungen mit Programmieren

Abfragen von Informationen oder Vorgehensweisen: ausgeprägt bei Gruppenarbeit von Erstsemestrigen -> Schneller Einstieg in Thema?

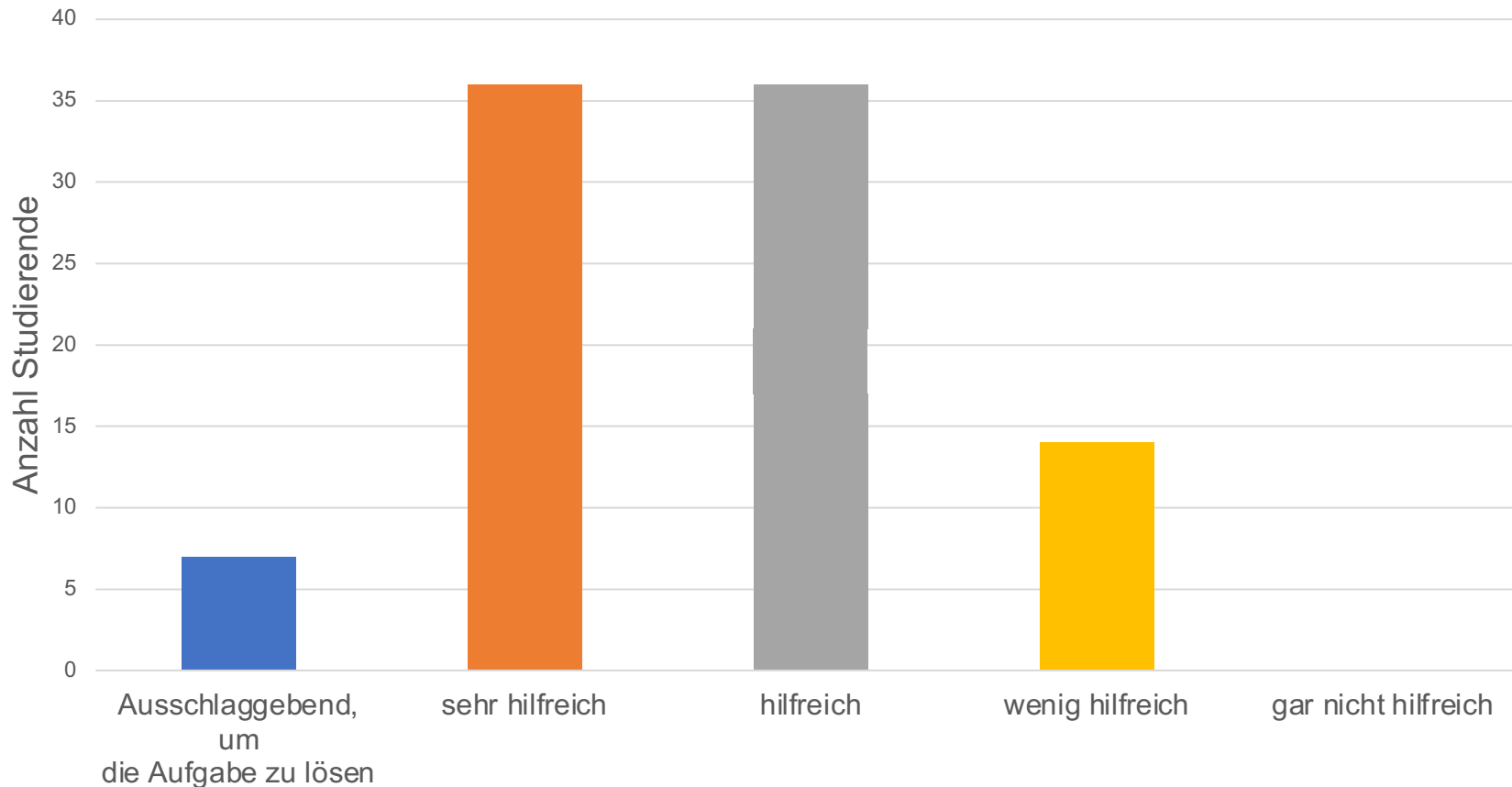
Analyse einer Aufgabe, Inspiration: Semesterarbeit (Konzeptarbeit), E-Prüfung (Modellerstellungen) -> Wissen der Studierenden im Thema? Advanced prompting nötig?

Faktor Zeit: Je mehr Zeit zur Verfügung stand, für desto mehr Tätigkeiten wurde ChatGPT verwendet

Wie schätzen Studierende die Nützlichkeit von ChatGPT bei den einzelnen Verwendungen ein?

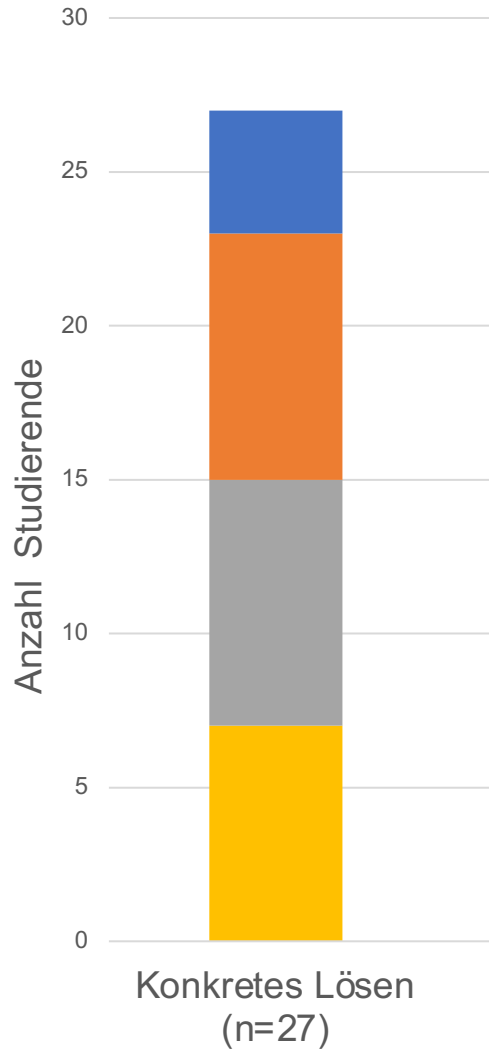
	Ausschlag ebend, um die Aufgabe zu lösen	sehr hilfreich	hilfreich	wenig hilfreich	gar nicht hilfreich	nicht verwendet
Analyse der Aufgabe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Abfragen von Informationen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inspiration	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Abfragen Vorgehensweise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie hilfreich war ChatGPT bei den Aufgaben insgesamt?



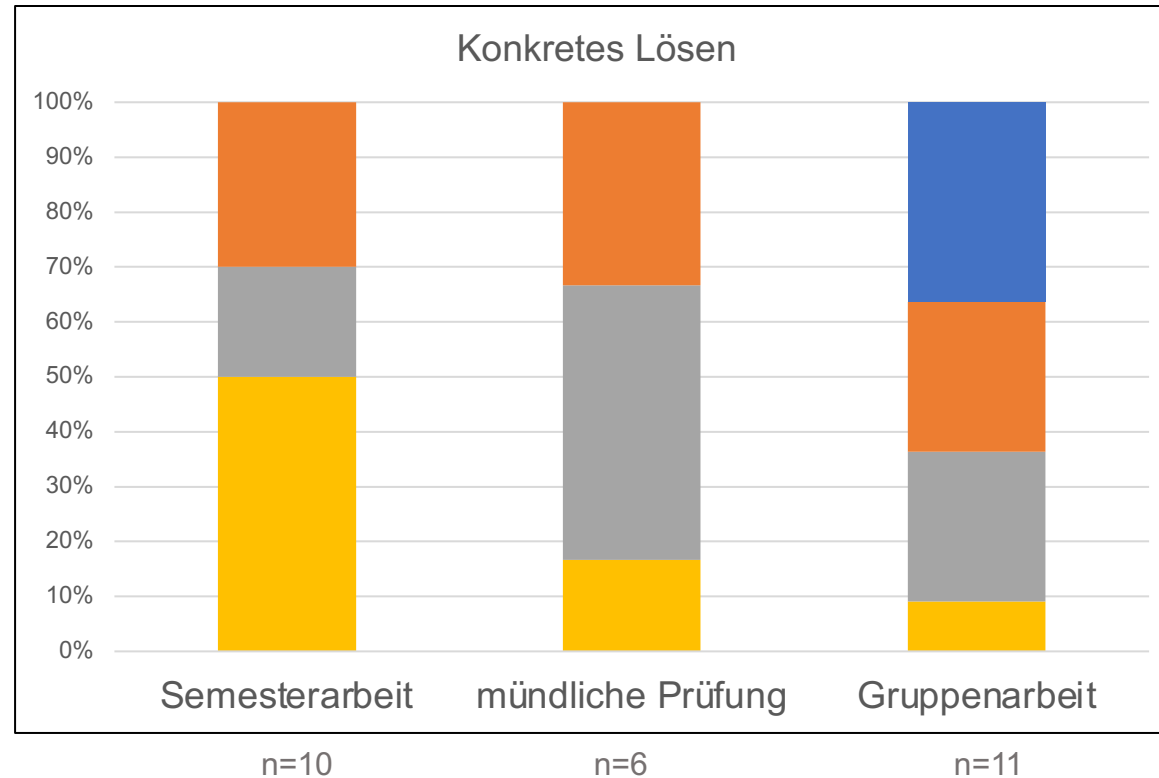
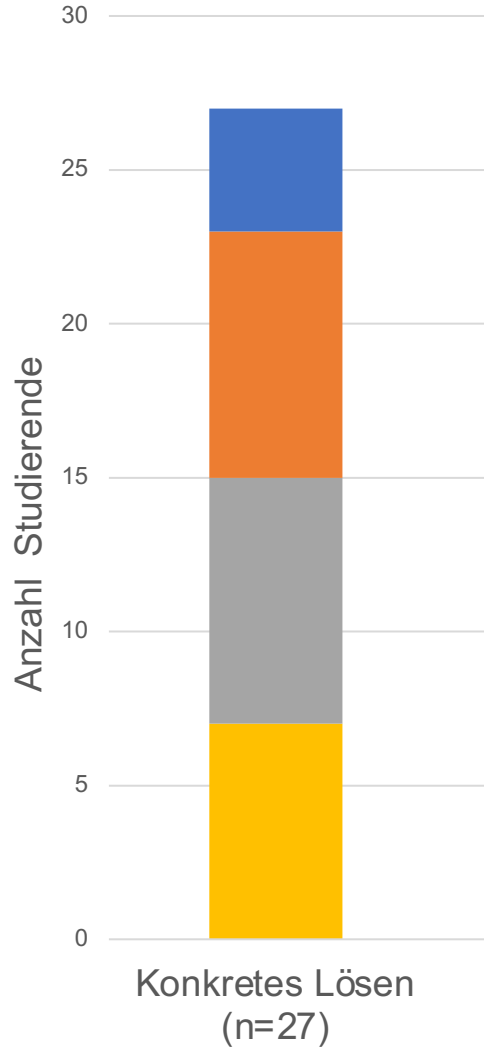
Wie hilfreich war ChatGPT beim konkreten Lösen?

Hauptverwendung, sehr ausgeprägt bei Prüfungen mit Programmieren



Wie hilfreich war ChatGPT beim konkreten Lösen?

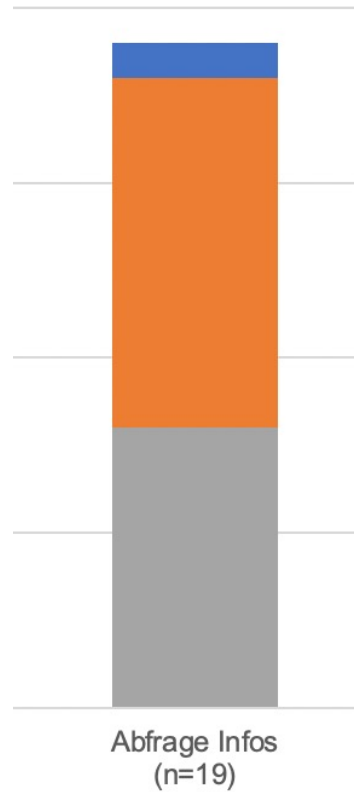
Hauptverwendung, sehr ausgeprägt bei Prüfungen mit Programmieren



- Ausschlaggebend, um die Aufgabe zu lösen
- sehr hilfreich
- hilfreich
- wenig hilfreich

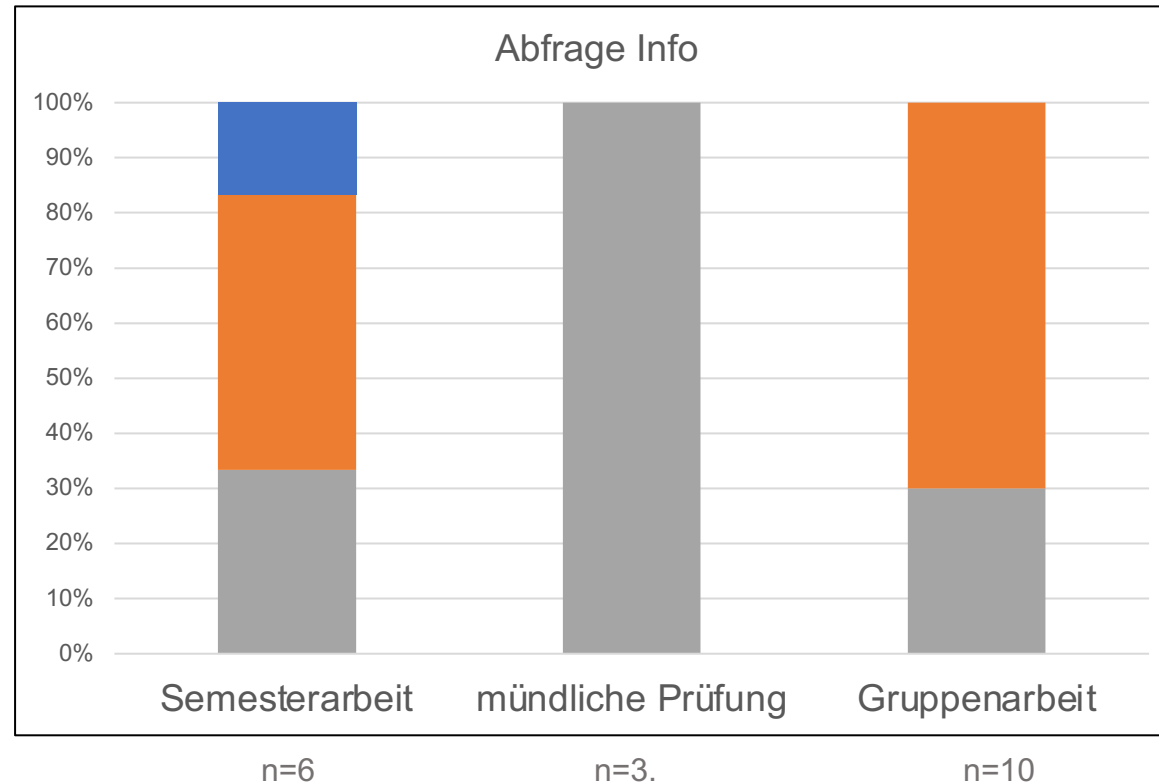
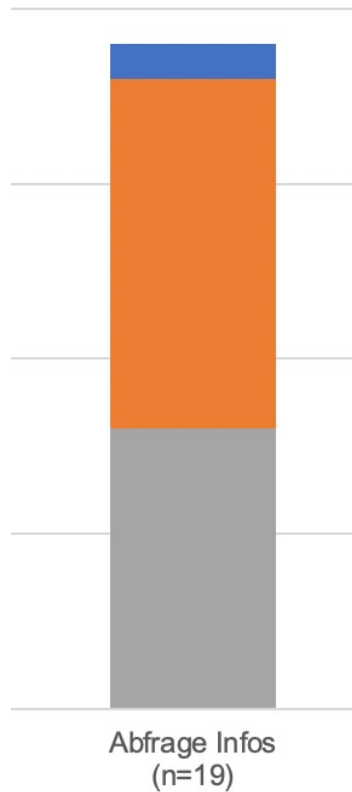
Wie hilfreich war ChatGPT zur Abfrage von Infos?

ausgeprägt bei Gruppenarbeit von Erstsemestrigen -> Schneller Einstieg in Thema?



Wie hilfreich war ChatGPT zur Abfrage von Infos?

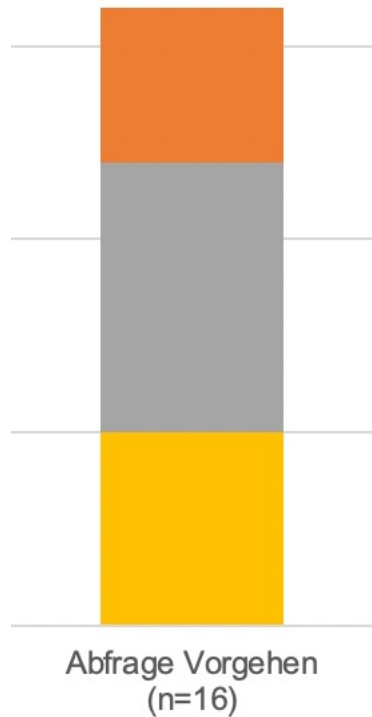
ausgeprägt bei Gruppenarbeit von Erstsemestrigen -> Schneller Einstieg in Thema?



- Ausschlaggebend, um die Aufgabe zu lösen
- sehr hilfreich
- hilfreich
- wenig hilfreich

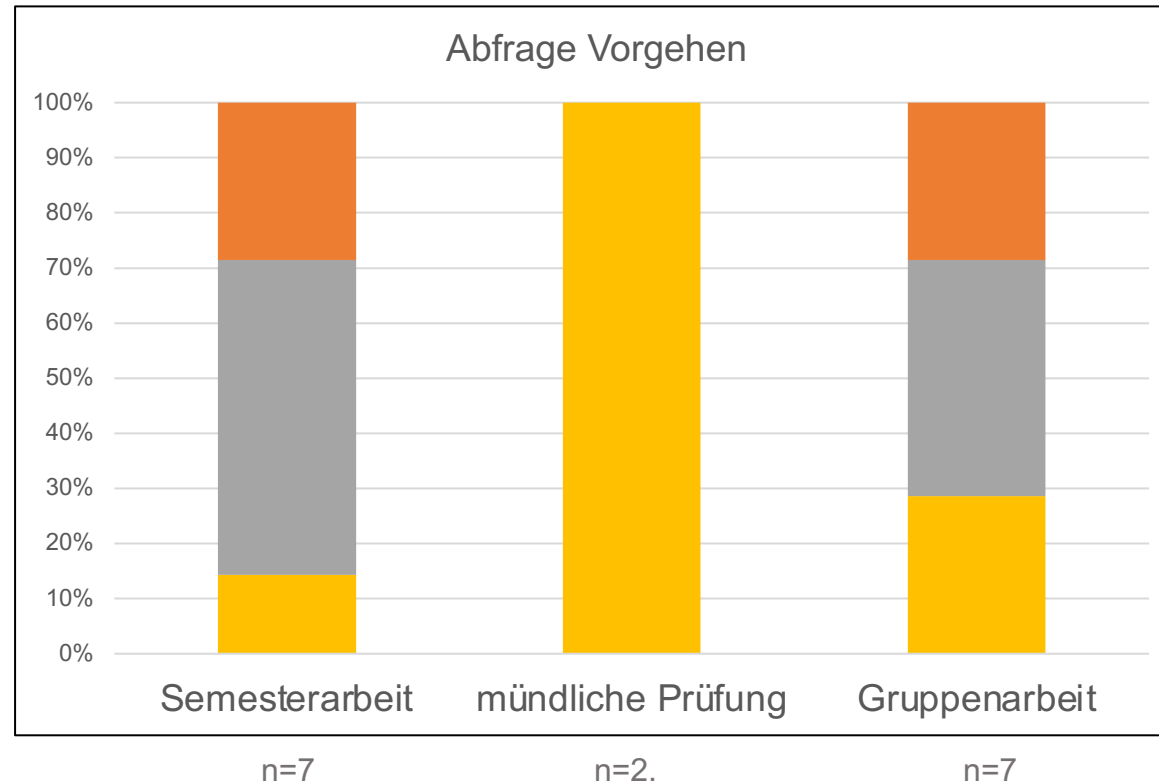
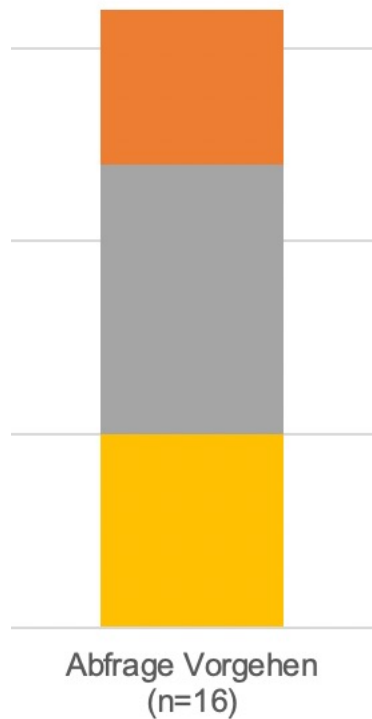
Wie hilfreich war ChatGPT zur Abfrage des Vorgehens?

ausgeprägt bei Gruppenarbeit von Erstsemestrigen -> Schneller Einstieg in Thema?



Wie hilfreich war ChatGPT zur Abfrage des Vorgehens?

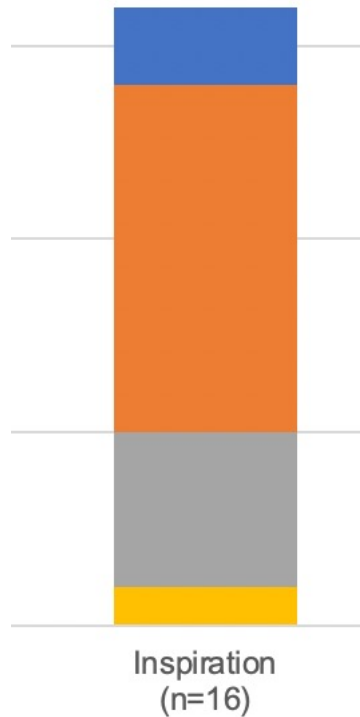
ausgeprägt bei Gruppenarbeit von Erstsemestrigen -> Schneller Einstieg in Thema?



- Ausschlaggebend, um die Aufgabe zu lösen
- sehr hilfreich
- hilfreich
- wenig hilfreich

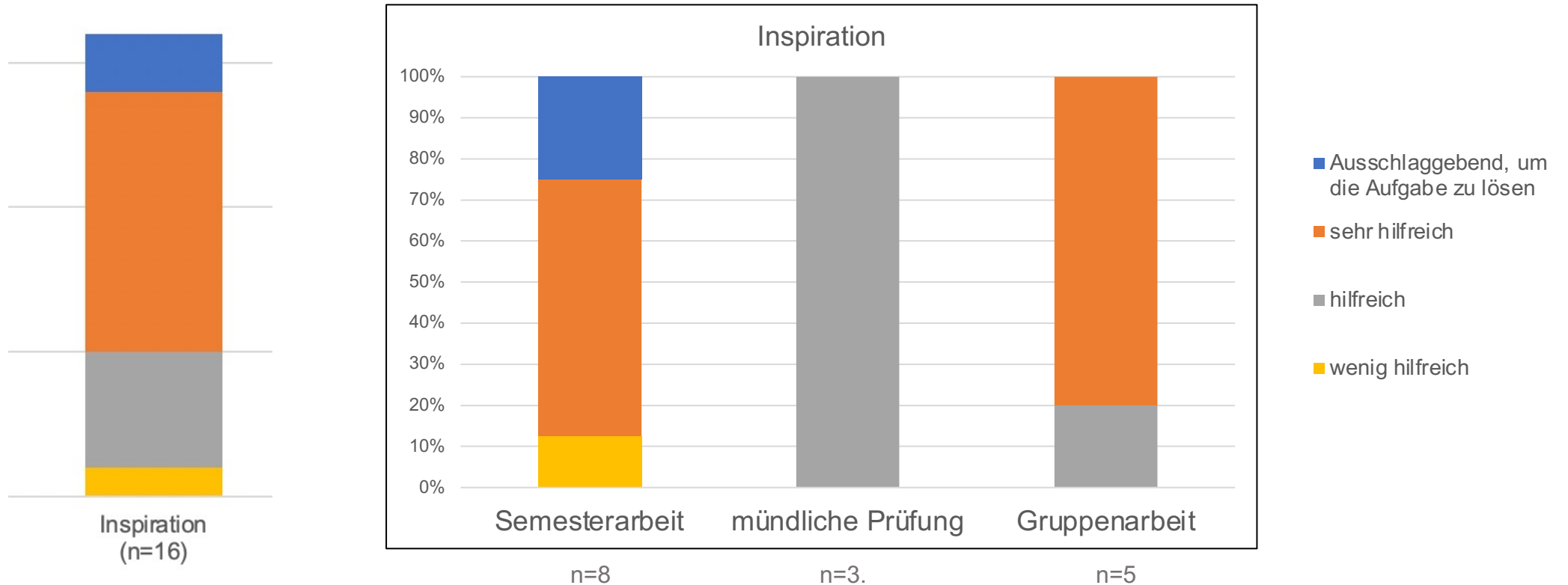
Wie hilfreich war ChatGPT zur Inspiration?

Semesterarbeit (Konzept, Modelle) -> Wissen der Studierenden im Thema? Advanced prompting nötig?



Wie hilfreich war ChatGPT zur Inspiration?

Semesterarbeit (Konzept, Modelle) -> Wissen der Studierenden im Thema? Advanced prompting nötig?

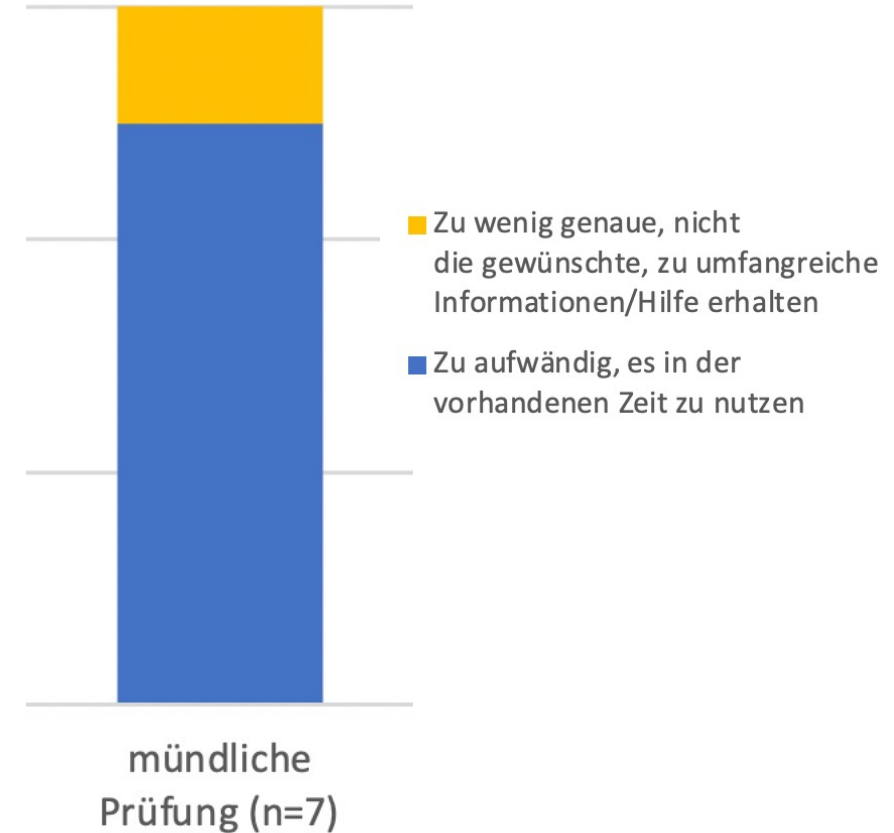


Welche Limitationen haben Studierenden erfahren?

Zu aufwändig, es in der vorhandenen Zeit zu nutzen

mit ein zwei mehr prompts hätte ich dies wohl auch hinbekommen, jedoch war der Druck innerhalb der mündlichen Prüfung eher gross, was dazu geführt hat, dass man sich nicht mehr die Zeit nehmen konnte und das Problem in 5 Minuten "sauber" zu lösen.

Die Wartezeit, die sowohl durch das Verfassen des Prompts als auch durch die darauf folgende Reaktionszeit entsteht, führt zu einer unangenehmen Situation der Stille.

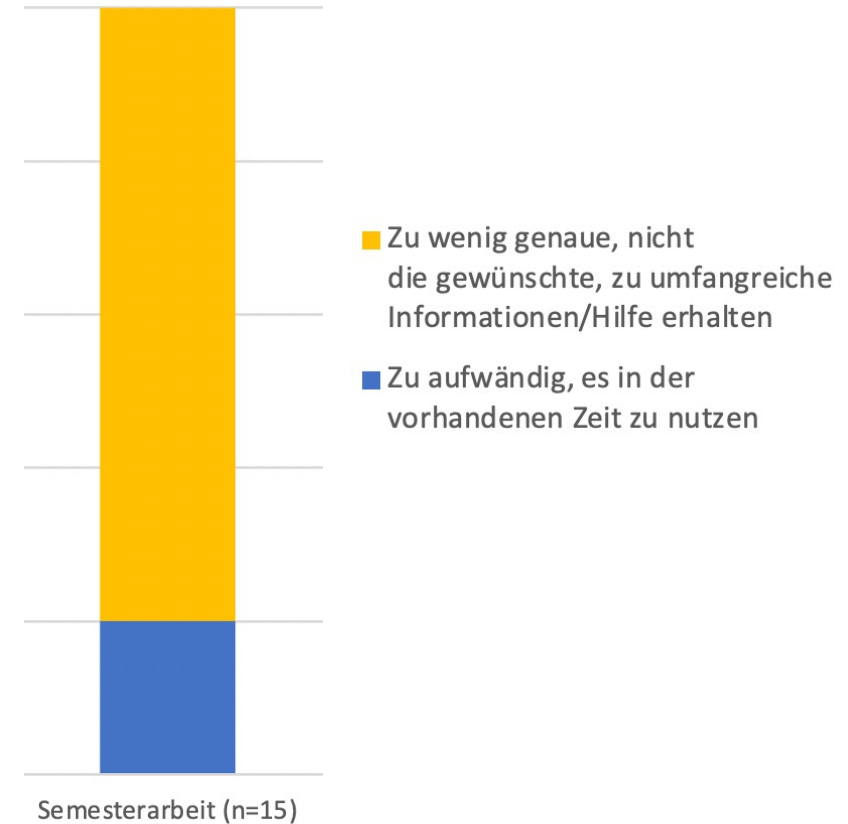


Welche Limitationen haben Studierenden erfahren?

Zu wenig genaue, nicht die gewünschte, zu umfangreiche Informationen/Hilfe erhalten

Zum Teil sehr allgemeine Antworten, welche immer spezifiziert werden mussten. Teils war man schneller, wenn man die Frage direkt selber löste ohne die Hilfe von ChatGPT.

Teils zu grosses Antwort Volumen mit zu wenig konkreten Inhalt und Wiederholungen.



Was würden Sie anderen Studierenden raten?

Prompting skills!

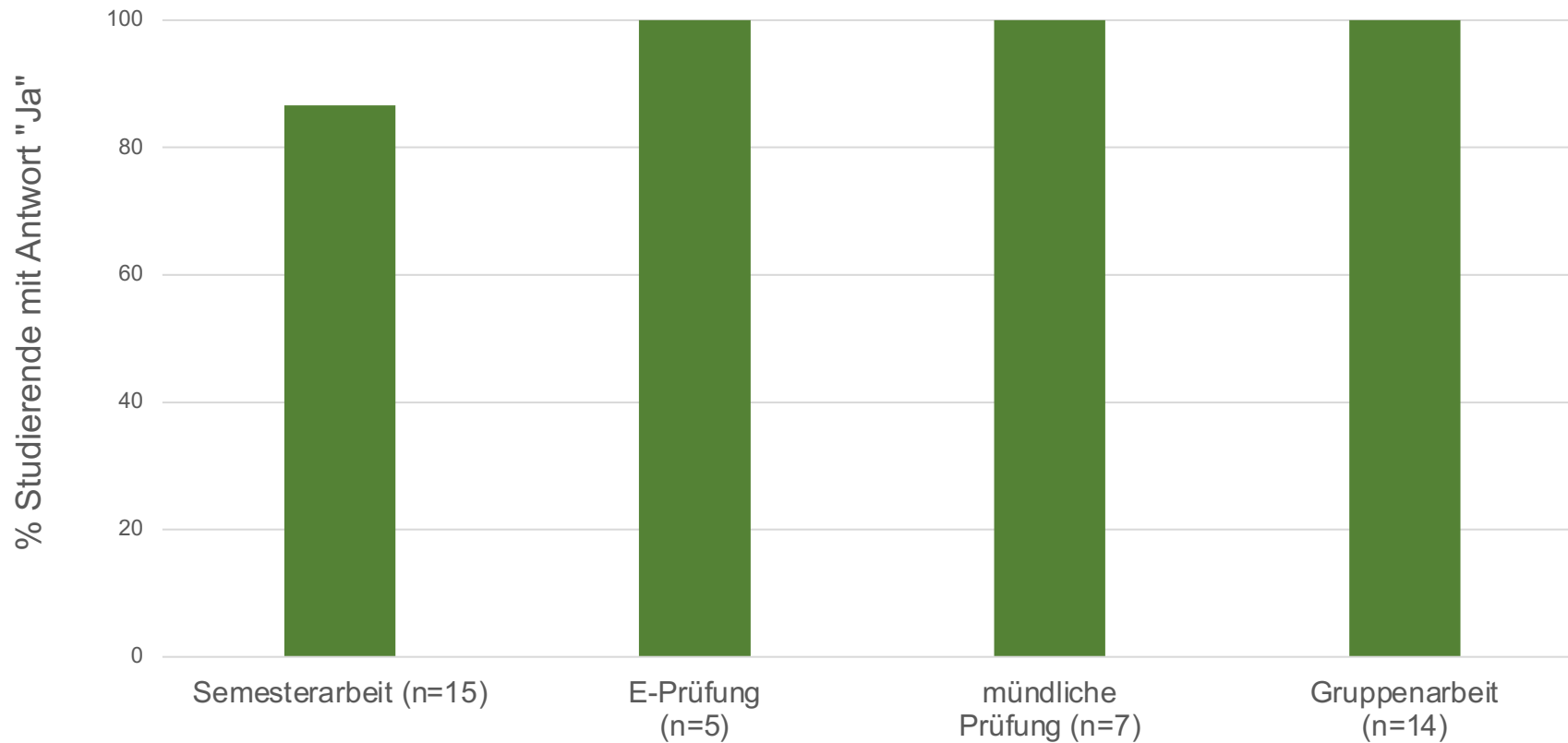
- Ausgangslage/Rahmenbedingungen genau schildern
- Aufgabestellung klar definieren
- Anweisungen und Fragen konkret und präzise formulieren
- Mehrere Male fragen
- Ergebnisse überprüfen

Was würden Sie anderen Studierenden raten?

Vorwissen haben!

- *Ist sehr hilfreich, um schnelle Informationen bei Planlosigkeit zu erhalten*
- *Man sollte vorsichtig sein, weil nicht alles richtig ist was im ChatGPT ist*
- *Vorkenntnisse sollten vorhanden sein, damit ChatGPT zielgerichtet kann genutzt werden (Aufwand/Ertrag). Sich auf ChatGPT verlassen um für fehlende Kenntnis von gefragten Konzepten zu kompensieren, ist nicht zielführend und hat bei mir nicht funktioniert*
- *Sicher gehen, dass die Outputs verstanden werden. Sonst ist es eher schädlich*

Würden Sie ChatGPT an dieser Prüfung wieder nutzen?



Fazit

Studierenden nutzen und empfehlen die Nutzung von ChatGPT. Trotzdem sind Verwendung & Nützlichkeit limitiert:

- Es gibt Aufgaben, bei denen sich die Verwendung von ChatGPT anbietet, bei genügend Zeit auch ausgiebig verwendet wird und sich als nützlich erweist
- Faktor Zeit = Stress
- Missverhältnis Aufwand & Ertrag = (noch) fehlende Kompetenz in Prompt Engineering

=> neues Instrument, dass man erst kennen- und gut zu nutzen lernen muss

- Ohne Fachkenntnisse geht es nicht (gut)

Was beobachten Dozierende bei der Verwendung von ChatGPT in ihren Prüfungssettings?

Einschätzung Dozierender Gruppenarbeit

Nützlich bei

Einfachem Programmieren, Code für
Visualisierungen, Debugging
-> grosse Zeitersparnis

Nicht nützlich bei

komplexeren Aufgaben
-> nicht effizient

geringer Kompetenzerwerb (“sind in der Aufgabe nicht weit gekommen”, “haben keine Verständnisfragen gestellt”)

Geringe Erfahrungen der Studierenden mit Prompting- es wurde rasch auf bekannte Mittel zurückgegriffen bzw die persönliche Unterstützung wurde bevorzugt

Einschätzung Dozierender Mündliche Prüfung

Nützlich weil

Bei diesem Modul hat ChatGPT eine grosse Informationsbasis

Nicht nützlich weil

Prompting braucht viel Zeit

Prompting gibt Hinweis auf Kompetenz der Studierenden
Viel Improvisation des Prüfers notwendig

Prüfung muss angepasst werden, zB Kriterien für Lösungen erarbeiten lassen anstatt Lösung selbst

Einschätzung Dozierender Semesterarbeit

Nützlich als

Hilfsmittel, Forschungsassistent

-> analog guter Suchmaschine

-> beschleunigt Rechercheprozess

Nicht nützlich bei

Fehlendem Fachwissen

Suche nach "pfannenfertigen" Lösungen

Wahl einer Vorgehensweise

Kreativprozessen

Programmieren mittels graphischer Oberfläche

Keine Verfälschung der Leistungen der Studierenden durch die Nutzung von ChatGPT festgestellt
Einsatz von KI könnte in Zukunft ein Zusatzkriterium werden

Einschätzung Dozierender E-Prüfung

Nützlich bei

Überprüfen von Code

Nicht nützlich bei

Lesen von Diagrammen (FS23)

Schreiben von sehr kurzen und präzisen

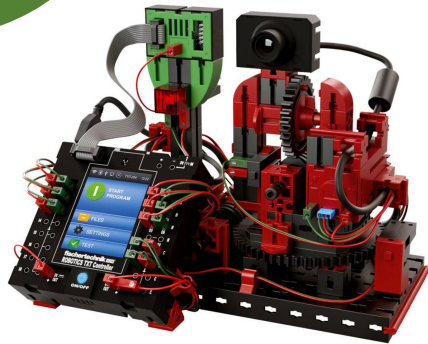
Fachtexten/Definitionen

Höheren Kompetenzstufen (Bewertung KI Output)

Eigenes Wissen und nicht Prompting Kompetenz prüfen

Was heisst das für meine momentane Prüfung?

Mindestens
ein Teil der
Prüfung



Wie prüfen ich
mein Lernziel?



nein

ja

Geeignete
kompetenzorientierte
Prüfung

Geeignete
kompetenzorientierte
Prüfung

Lernziel, das
auch künftig
ohne KI
beherrscht
werden muss?

```
import re  
hand = open('mbox-short.txt')  
for line in hand:  
    line = line.rstrip()  
    if re.search('From:', line):  
        print(line)
```

mit
Eigenständigkeits
erklärung und
Hilfsmittel
deklaration



Prüfungsformen

Open Book Prüfung

Semesterarbeiten

E-Portfolio

offene
Prüfungsformen

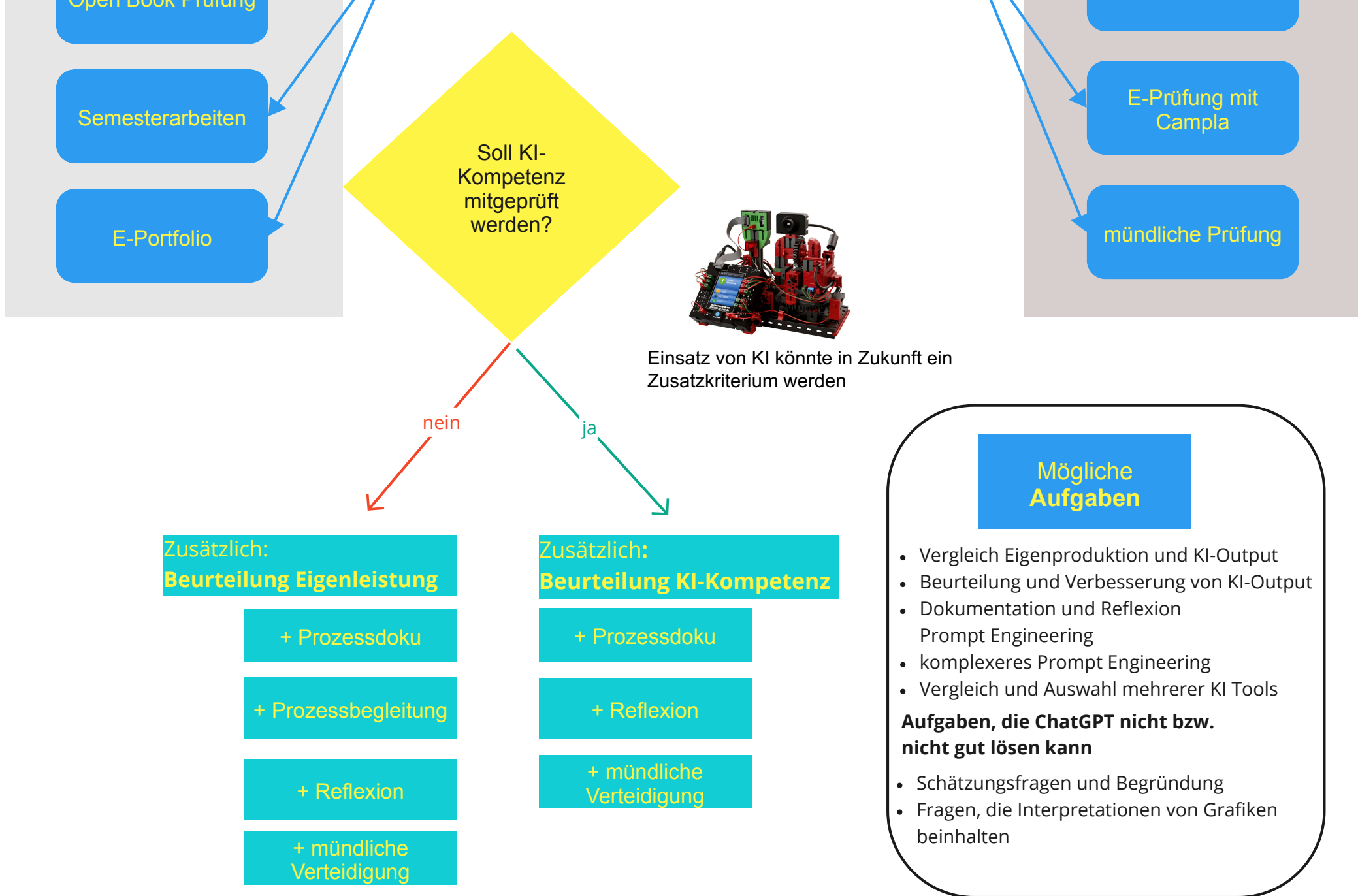
geschlossene
Prüfungsformen

Prüfungsformen

Paper & Pencil

E-Prüfung mit
Campi

mündliche Prüfung



Welche Schlüsse können wir weiter für die Gestaltung kommender Prüfungen ziehen?

Um den Nutzen von ChatGPT im eigenen Modul/in der eigenen Prüfung abzuschätzen, muss **auch selbst ausprobiert werden, was ChatGPT kann oder eben nicht kann.**

Um alle Lernziele prüfen zu können, **mehrteilige Prüfungen** erstellen: Geschlossener Teil für Wissensabfragen, offener Teil für realitätsnahe, kompetenzorientierte Aufgaben (Bsp: OOP in SGI), zusätzlicher mündlicher Teil

Es darf nur geprüft und bewertet werden, was den Studierenden kommuniziert wurde und was sie üben konnten! (Lernziele + Lernmöglichkeiten = Prüfung)

Soll die **Verwendung von KI** zugelassen werden, müssen die Studierenden bei der Nutzung von KI im Modul unterstützt werden (zB Prompting Skills erwerben)

Will man **auf höheren Bloom Stufen prüfen** (zB problemorientierte Aufgaben, bei denen zuerst eine Analyse erfolgen muss; Kriterien für gute Lösungen entwickeln lassen anstatt direkt eine Aufgabe zu lösen), müssen die Lernziele und Übungsmöglichkeiten entsprechend angepasst werden

In jedem Falle müssen auch die Bewertungskriterien angepasst werden.

Fazit

Wir werden nicht darum herumkommen, uns selbst mit generativer KI zu beschäftigen und unsere Module (Lernziele, Lerninhalte, Lernmöglichkeiten, Prüfungen) entsprechend anzupassen