Interview with S5 and S6 - Anonymized German transcript of the statements  
 with reference to NodeGrade

*(S = Student; I = Interviewer)*

I: *Gibt es konkrete Verbesserungen, die euch noch einfallen und ihr euch wünschen würdet?*

S5: Freitext bei der Aufgabe, da sind wir auch lang hängen geblieben, weil es ist gut, was selber zu formulieren, weil man muss sich dann Gedanken darüber machen und wie formuliere ich das und das ist dann cool, dass man auch ein Feedback bekommt, wo es dann heißt, das ist gut, das ist vielleicht schlecht, aber wir haben Feedback bekommen von wegen, das ist schlecht und wir so, okay, wir schreiben das jetzt dazu und schicken es nochmal ab und dann war es immer noch nicht perfekt, weil dann plötzlich was Neues aufgefallen ist und dann sind wir so ein bisschen an dem Punkt gewesen, so okay, müssen wir jetzt den Ding so schrittweise immer, das müssen wir jetzt schrittweise immer verbessern und dann kriegen wir so nach und nach die Fehler alle aufgelistet, weil es hat halt eine Sache explizit als falsch oder als fehlend angegeben, wir so, okay, dann muss ja das andere passen, also gut, dann fügen wir das ein und dann war es immer noch nicht okay und es war so ein bisschen…

Bei der Freitextaufgabe sind wir lange hängen geblieben. Es ist gut, selbst etwas zu formulieren, da man sich Gedanken machen muss, wie man es am besten formuliert. Es ist dann cool, dafür ein Feedback zu erhalten, was vielleicht gut und was vielleicht schlecht ist. Aber wir haben Feedback bekommen im Sinne von „XY fehlt“. Dann haben wir das dazugeschrieben und noch abgeschickt, doch es war immer noch nicht perfekt, weil etwas Neues aufgefallen ist. Wir sind dann an dem Punkt gewesen, okay, müssen wir das jetzt schrittweise verbessern und kriegen nach und nach die einzelnen expliziten Fehler alle aufgelistet? Da jeweils nur einzelne Aspekte als falsch oder fehlend markiert wurden, sind wir davon ausgegangen, dass der Rest passt. Wir haben also nur den Fehler korrigiert, aber dann war es immer noch nicht okay…

S6: Das ist ja auch ein sehr typisches AI-Problem. Wenn ich jetzt GPT sage, schreibt mir mal den Code und dann sage, der Code ist nicht gut, dann schreibt es einen neuen. Wenn ich dann wieder sage der ist auch nicht gut, schreibt es wieder einen neuen. Aber im Endeffekt kriegst du immer eine andere Antwort, die gleich richtig ist. Und bei uns war es jetzt zum Beispiel so, dass wir haben beide das Fachwort benutzt haben und dann haben wir gesagt, ja, aber das müssen wir nicht direkt reinschreiben, weil das ergibt sich aus dem Kontext. Ich finde alles mit Fachwörtern zu beantworten nicht immer sinnvoll, wenn man nur sein Verständnis zeigen will. Als wir es dann aber nachgefügt haben, war es nicht konkret genug. Die Frage war aber nicht, „nenne dann Beispiele dazu“ oder so. Die Sache ist aber, wenn ich jetzt Beispiele nennen soll, dann mache ich was Konkretes. Aber wenn die Frage einfach nur ist, „Was ist der Vorteil?“, dann fange ich nicht an einen Roman nur über Beispiele zu schreiben im Sinne von „hier ist der Beweis und das ist so und so“.

S5: Die Aufgabenstellung war sehr allgemein. Wenn da jetzt gestanden wäre „in Bezug auf Beispiel XY“ oder „bring ein Beispiel XY und beziehe dich drauf“, dann hätte ich das verstanden mit dem spezifizieren aber die Frage war halt recht allgemein. Wir haben dann allgemeiner geantwortet und das Ding so, „ja, geht da spezifischer drauf ein“. Ich glaube, was wir eingefügt haben, war, dass das unerwartete Verhalten auftritt. Da haben wir gesagt, das ergibt es sich aus dem Kontext. Die Antwort war so, schreibt das konkret rein. Wir dachten uns, okay dann schreiben wir es konkret rein. Es wollte dann aber wieder, dass wir spezifisch drauf eingehen. Laut Aufgabenstellung sollen wir aber doch keine Beispiele nennen. Das war dann so ein bisschen komisch...

S6: Da muss man vielleicht gucken, dass wenigstens die Rückantwort der Aufgabenstellung inhaltlich näher kommt. Sonst war es eigentlich okay. Es ist halt, wie gesagt, diese typische automatische Rückantwort. Das hätten wir perfektionieren können, bis wir alt sind.

S5: Ich bin bei sowas sehr perfektionistisch, weil ich halt in der Klausur das ja logischerweise auch so gut machen will, wie es geht und wenn das Ding mir dann halt sagt, ich habe 76 Prozent oder Punkte, wie auch immer, dann ist das so…

S6: Würde mir auch nicht reichen.

S5: Da könnte mehr sein. Aber wie komme ich zu diesen 100%? Er sagt mir dann halt, „verbessere das“. Ich verbessere das und dann habe ich immer noch nicht die 100%.

S6: Wenn man vielleicht 80 erreicht, sind das für das Tool dann 100. Ist jetzt doof, aber irgendwo muss diese Grenze sein, dass man sich sicher ist, ich habe das richtig beantwortet. Ich zum Beispiel wäre da sonst ewig dran um es zu perfektionieren.

S5: Mir ist tatsächlich noch aufgefallen: die Musterlösung muss man irgendwo dann noch einsehen können. Weil dann wäre uns, glaube ich, aufgefallen, was speziell noch fehlt.

S6: Das wäre gut ja.

I: *Okay, jetzt für die Rückmeldung zur Antwort.* *Wie nützlich fandet ihr den Score, diese Klassifikation und die schriftliche Fassung? Wenn man das einmal einzeln betrachtet: Wir haben ja einmal diese Zahl gehabt, dann gab es diesen Score mit „bestanden“, „nicht bestanden“, beziehungsweise die Färbung der Leiste und das schriftliche Feedback. Wie schätzt ihr die Nützlichkeit der einzelnen Sachen ein?*

S5: Manchmal war ich ein bisschen verwirrt, ab wann es jetzt bestanden ist. Also ich hätte da gerne eine Definition, weil ich bin da sehr perfektionistisch bei sowas und würde halt gerne wissen, okay, was heißt jetzt „Bestanden“? Ich habe irgendwie sowas wie 80% oder so, heißt es, ich habe alles richtig? Eventuell in dem Balken eine Linie ab wann man bestanden hat. Das hätte mir schon gereicht.

S6: Eigentlich war das bei der Frage nur „Number Go Bigger“. Aber ich weiß nicht, was ich dazu formulieren soll. Bis auf die Ausnahmen, die wir jetzt gesagt haben, bin ich sehr zufrieden gewesen. Das mit der Musterlösung ist, wie gesagt, eine super Idee, weil dann hat man wirklich Orientierung. Aber es ist auch die Frage, woran ist die Aufgabe angelehnt? Ist das eine Klausuraufgabe? Weil dann würde ich keinen Roman schreiben, weil ich die Zeit dafür nicht habe. Ist es aber einfach so eine Übungsaufgabe für zu Hause, dann könnte man auch ruhig einen Roman schreiben. Dafür hat man die Zeit, technically. Aber wenn ich sag, boah, ich muss die jetzt unter zwei Minuten lösen, dann spezifiziert man sich nicht so. Also da ist halt die Frage, was man halt haben will, vielleicht die Erwartungen so ein bisschen kommunizieren. Ich glaube, es sind größtenteils eigentlich nur die Erwartungen, die da so ein bisschen gefehlt haben. So die Enderwartung, was ist wirklich die Lösung und wie doll soll ich denn beantworten? Also sind wirklich die Erwartungen am Ende.

S5: Ja, bei einer Klausur heißt es ja oft so man hat irgendwie 60 Minuten, es gibt 60 Punkte, also quasi einen Punkt pro Minute, sodass man so ungefähr weiß, okay, wie zügig muss ich jetzt was beantworten. Wenn es zwei Punkte gibt, dann habe ich quasi einen Aufwand von zwei Minuten in dieser Aufgabe und da kann man es irgendwie so ein bisschen einstufen. Bei der Aufgabe habe ich jetzt, quasi kostennutzenmäßig, gar nicht die Zeit einen Roman zu schreiben, weil es gibt nur zwei Punkte.

S6: Dann kann man bei der Aufgabe schreiben, nehme dir etwa zwei Minuten Zeit.

S5: Ja, dass man ein irgendwie ein bisschen die Gewichtung reinbringt. Weil ich kann jetzt eine Seite schreiben und der kann mir dann immer noch sagen, „oh, das ist jetzt zu unspezifisch“. Wenn ich aber in der Klausur bin, kann es sein, dass diese Seite schon viel zu viel ist.

I: *Dann ist noch die Frage, hat das Tool euren Lernprozess unterstützt und würdet ihr mehr oder andere Aufgaben noch mit diesem Tool wünschen? Und wenn ja, welche? Wie würde das ausschauen für euch?*

S6: Erst habe ich so gesagt, ich hasse Texte, ich will das nicht beantworten. Aber als ich es dann einmal gemacht habe, fand ich es dann gar nicht so schlecht. Also es ist so ein bisschen so: Wenn man einmal angefangen hat zu schreiben, dann kann man auch mal gucken, ob seine eigene Denkweise darüber richtig ist. Normalerweise kann man immer nur Sachen abhaken. Aber jetzt kann ich halt sagen, ich denke da aber so drüber. Vielleicht ist die Frage, ob man das dann selbst erkennt. Was ist, wenn ich es jetzt nicht verstanden hätte? Wenn ich es jetzt wirklich falsch verstanden hätte? Es hätte mir jetzt gesagt, es geht mit dem Hund spazieren und ich hätte ihm erzählt, „ja, die Katze führt man morgens um acht aus“. Hätte er mir dann vernünftig gesagt, „ey, wir reden über das Hundespazieren, dir fehlt das, das und das“? Soweit haben wir es ja nicht getestet. Wenn wir es jetzt mal negativ getestet hätten, hätte mich mal interessiert, ob er mich auf den richtigen Weg gebracht hätte. Ich hätte auch gerne mehr Aufgaben. Auch jetzt in unserem Test hätte ich schon gerne eine pro Thema mindestens gehabt, aber ich verstehe, dass das für den Test nicht mehr möglich war.

S5: Ich finde die auch an sich sinnvoll, nur muss man das irgendwie einstufen können und das Feedback muss irgendwie greifbar sein, was es jetzt konkret von einem will.