Naïve concepts of aerodynamic lift – data lessons from different (learning) cultures

Florian Genz
*Future Strategy for Teacher Education (ZuS), Universität zu Köln, Germany*

Kathleen Ann Falconer
*Institute of Physics Education, Universität zu Köln, Germany*

Inclusivity is a premise for successful scientific discourse in the endeavor of gaining a broader understanding. In this work therefore, we elicit different “schools of thought” in the context of flight physics for reducing unconsciously exclusionary practices. For decades aerodynamic lift explanations have been a highly controversial topic in PER – and they still are. However, the discussion has always been mainly driven by disjunct models and hermeneutical arguments. This paper tries to carve out empirically different “schools of thoughts” by asking 400+ university students at different institutions about their agreement to various explanations for aerodynamic lift. The study was accompanied by the expert-validated Flight Physics Concept Inventory (FliP-CoIn) and therefore can differentiate what high performers, low performers and flight instructors think. The results surprised even the authors: Among many other findings, this study revealed that – within ONE mind – naïve concepts can coexist with expert concepts and that this phenomenon is especially prevalent among high scoring individuals of the FliP-CoIn instrument.
I. INTRODUCTION

Inclusion of diverse thoughts is a premise for successful discourse [1], but in their search for identity and belonging, human beings cluster into different social groups [2]. In and within these social groups learning cultures or schools of thought emerge, implicitly excluding ways to think. Eliciting these clusters can reduce unconsciously exclusionary practices and lead to a wider scientific discourse [3]. In this work, therefore, we will look at different naïve and expert conceptualizations in the context of flight physics. Furthermore, we will investigate the coexistence of different concepts [4] directly, which is a reported research gap [5].

Naïve concepts were found to be a key element to teaching and learning [6–8]. Therefore, science education and especially PER has established an extensive body of research about them for many contexts (i.e. on physport.org). However, concerning the concept of aerodynamic lift for fixed wings, there is relatively little consensus, even among experts [9]. Since aerodynamic lift is a complex phenomenon, different schools of thought have developed differing approaches to explain lift [10–17]. These are motivated by different scopes of application, simplification, and calculation methods. The models do not necessarily contradict each other, but for the novice learner, this coexistence of different expert concepts can be highly confusing. The limitations and applicability of a model is the hardest to grasp [18]. Therefore, it should not come as a surprise when students merge aspects from different models in their minds – even though these models are incompatible/disjunct to each other. We call this the model merging phenomena (MMP) (see FIG. 1).

We call this the mental model merge effect (MMME). It will be elaborated in the context of flight physics by the example of concepts for aerodynamic lift. Moreover, this work will shed some empirical light on the different schools of thought by asking students and flight instructors at different institutions about their agreement to various rationales for aerodynamic lift.

II. METHODOLOGY

Agreement to different explanations for aerodynamic lift was collected on a 4-point Likert scale for seven rationales (see also FIG. 5). Rationale 1 and 7 (R#1 & R#7) focus on different aspects of the naïve concept of path length reasoning to explain aerodynamic lift [19]. R#2 was designed to be agreeable for students with the skipping stone concept (=no air viscosity) [20]. R#3 is rather vague and should be agreeable for all schools of thought, but especially for those where the teaching focus is on downwash and/or smoke trail deflection. R#4 is about the pressure and might touch the Bernoulli principle [21] in the reader’s mind but without making the mistake to speculate about reasons for the observable velocity difference. R#5 focuses on the Coandă effect, boundary layer and near field [22] and R#6 on the circulatory flow and the Kutta condition [23]. All rationales were derived from literature [10–17, 24] and student or expert interviews [25]. The bilingual development of the questions was accompanied by language and physics education research experts [26].

A. Surveyed population

The survey consisted of seven rationale ratings (see also FIG. 5), demographic questions and the Flight Physics Concept Inventory (FliP-CoIn) [27], which allowed the authors to differentiate for different subgroups (e.g. students...
vs. flight instructors, high- vs. low scorers, etc.). We collected n=400+ complete online and paper-pencil questionnaires at three established tertiary education institutions for engineering and aviation between 2017 and 2018 (see FIG. 3):

- Dataset 1 (DS1) was collected as an online survey for all semesters with an inducement of a little extra course credit.
- DS2 was an online survey for advanced undergraduate semesters – for some courses during lecture time. Inducement was a raffle entrance for free bike rentals.
- DS3 was conducted via paper and pencil during the first session of an aerodynamics I lecture for early undergraduates. 71% of DS1 participants reported more than 25 hours or motorized flight, DS2 and DS3 had none.

For further analyses, all three datasets were divided into high and low scorers with the help of the Flight Physics Concept Inventory (FlIP-Coln) total score [27]. Following Kelley [28], the top-scoring 27% of each dataset were considered high scorers and the bottom 27% low scorers. All total score distribution shapes look similar to each other (using Kolmogorov-Smirnov tests with appropriate adjustments).

Additionally, tentative claims were warranted using think-aloud protocol with airfoil shapes and an alternative population.

### III. RESULTS & DISCUSSION

A first valuable insight into student thinking was led by the question: Do high scorers show different overall response patterns than low scorers – for the seven rationales for aerodynamic lift? (see FIG. 4):

![FIG. 4. Cumulative usage of answer options: By high and low scorers (The 27% of top and bottom scorers of each dataset)](image)

Only in dataset 2, high and low scorers show similar answer option usage. Students from dataset 2 were also the most (aerodynamically) specialized and advanced students in their study program. This homogeneity could be due to a leveling effect of the learning intervention and/or the drop-out rate towards the end of the semester. In both cases it can result in a school of thought over a longer period of time [40].

Looking at FIG. 4, there seems to be a slight preference among high scorers towards answer option 4 (= “completely agree”). This could indicate that high scorers are more confident in their answers in general. But how do we explain that high scorers tend to agree more to all rationales at the same time? We speculate that high scorers scan the answer options for familiar patterns/facts only and overread contradictory aspects.

Moving on to response patterns for each specific rationale, the results from FIG. 5 show that rationale #1 (R#1) has relatively high overall agreement. This tendency

### FIG. 5. Rationales for aerodynamic lift: Agreement on a 4-point Likert scale. Grey cells = visible to the participants.

![FIG. 5. Rationales for aerodynamic lift: Agreement on a 4-point Likert scale. Grey cells = visible to the participants.](image)
holds even true for the four surveyed certified flight instructors (CFIs) in DS1. However, we would like to highlight, that R#1 (path length existence) is only a premise for R#7 (equal transit time), which is a well-documented misconception [9]! Even angled barn doors without any curvature or camber will fly – and they have literally zero path length difference [29]! Our preliminary interpretation of this finding is: Students mistake the help construct path length for a real thing, which supports our conceptualizations around FIG. 1 and FIG. 2.

The low agreement to R#2 (skipping stone) might be due to the fact that students – unnecessarily – associate this rationale simultaneously with the often accompanying aspects of crossing particle paths and frictionless fluid misconceptions [30, 31]. We need to emphasize here that for a finite wing in three dimensions, there is yet another known way for air to transfer impulse to the wing than air particles colliding with the surface. So, at least, the 33% who completely disagree with R#2 in the given wording (see FIG. 5) may trouble educators. However, some later think-aloud interviews indicate that students thought of R#2 and R#3 (downwash) as opposed to each other and that complete agreement to R#3 would demand a complete disagreement for R#2 (skipping stone). The assumption of exclusivity of only one true theory [39, 40] might also be the reason for the high percentage of complete rejection for rationale #6 (36%).

One student utterance during the think-aloud interviews suggests, that the might confuse “circulatory flow” with “wingtip vortices”.

Differentiating by datasets as well as high and low scorers brought further insights (see FIG. 6): R#4 (pure pressure & velocity) has relatively high agreement in all datasets, but in DS1 and DS3, agreement to naïve R#1 (path length existence) is even slightly higher. This gives rise to the conclusion that only DS2 students learn about the little but important nuances between R#4 (pure pressure & velocity) and R#1 (path length existence). Furthermore, DS2 is the only one where high scorers vastly disagree with R#7 (equal transit time).

For R#5 (Coanda effect), it is insightful to look at the tendency to the middle, because no expert asked so far (n=8) answered with “completely agree” or “completely disagree”. From low to high scorers there is a clear trend towards the two middle answer options in DS1 (54%→69%) and DS2 (50%→62%) whereas in DS3 there is a clear trend away from the middle answers (65%→44%).

For R#7 (equal transit time), the huge shift towards disagreement between low to high scorers in DS2 is most obvious. This effect does not occur in the other DSs. In DS3 the most frequent answer even worsened from “mostly agree” (50% for low scorers) to “completely agree” (47% for high scorers). This all supports the finding that high scorers in particular are at risk to – also – agree to naïve concepts and may even gain confidence in them [32-34]. This effect may foster the emergence of an increasingly dogmatic school of thought or, more fashionably called, a filter bubble. For more results and data representation we refer to the accompanying poster with identical title.

IV. CONCLUSIONS

A. Model limitations are key

Since rationale #1 (see FIG. 5) is a common misconception, it was not surprising to see that the agreement is very high in all datasets (DSs). What is surprising is that, in DS1 and DS3, the high scorer (top 27% of participants) agreement to this naïve concept is even higher than the low scorer (bottom 27% of participants) agreement.

<table>
<thead>
<tr>
<th>Rationale #:</th>
<th>path length existence</th>
<th>skipping stone</th>
<th>downwash</th>
<th>pressure &amp; velocity</th>
<th>Coanda effect</th>
<th>circularity flow</th>
<th>equal transit time</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>10%</td>
<td>35%</td>
<td>55%</td>
<td>15%</td>
<td>60%</td>
<td>25%</td>
<td>55%</td>
</tr>
<tr>
<td>#2</td>
<td>25%</td>
<td>25%</td>
<td>50%</td>
<td>40%</td>
<td>50%</td>
<td>30%</td>
<td>45%</td>
</tr>
<tr>
<td>#3</td>
<td>30%</td>
<td>40%</td>
<td>70%</td>
<td>35%</td>
<td>65%</td>
<td>35%</td>
<td>65%</td>
</tr>
<tr>
<td>#4</td>
<td>50%</td>
<td>50%</td>
<td>100%</td>
<td>60%</td>
<td>100%</td>
<td>75%</td>
<td>100%</td>
</tr>
<tr>
<td>#5</td>
<td>80%</td>
<td>80%</td>
<td>100%</td>
<td>75%</td>
<td>100%</td>
<td>85%</td>
<td>100%</td>
</tr>
<tr>
<td>#6</td>
<td>90%</td>
<td>90%</td>
<td>100%</td>
<td>85%</td>
<td>100%</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>#7</td>
<td>95%</td>
<td>95%</td>
<td>100%</td>
<td>90%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

FIG. 6. Agreement to rationales for aerodynamic lift: by dataset and high/low scorers (>27% of top and bottom scorers).
Moreover, in DS1 zero of 49 high scorers marked the option "completely disagree" and 30 of 49 "completely agree". This gives rise to overlapping hypotheses to be tested in further research: A) current learning interventions accelerate naïve conceptions B) naïve concepts are not contrasted enough with expert concepts C) students are not given enough opportunity to discuss and adapt their own mental models for grasping the limitations and applicability of models.

B. Concepts continue to coexist or stay merged

Rationale #7 focuses on a different aspect of the naïve pathlength reasoning (rejoining of air packages and equal transit time). Compared to rationale #1 (path length existence), the overall agreement scores are slightly lower but what stunned us was that only in DS2 the high scorers showed a much higher disagreement to rationale #7 (73% of high scorers in DS2 answered with "completely disagree" or "some agreeable facts" whereas only 36% of low scorers in DS2 marked one of these two options), whereas DS1 shows little differences between low scorers and high scorers and DS3 even shows a trend towards complete agreement. This gives rise to the idea, that at least at the institution of DS2 the "air packets rejoining" aspect of the pathlength misconception is well contrasted and debunked. However, naïve rationale #1 is still seductive in all datasets – and even more for high scorers (complete disagreement is always lower). Therefore, we conclude that debunking one aspect of a misconception (rejoining of air packages) might not be enough for students to drop it completely (=path length). The data suggests that they can exist independently from each other. A recent fMRI study seems to back this concept coexistence hypothesis [5]. However, it might also be the case that the model merge phenomena (MMP) (see. FIG. 1 & FIG. 2) might become a more practical approach for relating what happens during conceptual learning. Further studies should be able to differentiate that.

C. What do these findings mean for educational practice?

For the context of aerodynamic lift we showed that – especially in high scorer minds – different naïve concepts can continue to coexist next to expert concepts. Therefore, we recommend to shift educational effort away from replacing naïve concepts by expert concepts (usually by readings, lecture, contrasting misconceptions in theory,...) but rather let students actively find the limitations and strengths of their current concepts. This may be best facilitated by the help of simulations, experiments, concept mapping [6], real world observations and authentic, practical problems, as well as participation in scientific discussion. As impressively demonstrated by Derek Muller, we have good reason to believe that every direct/explanatory teaching effort is doing more harm than good [32]. Direct teaching efforts steal students time for first-hand experience and for own mental effort. As a result, students only become more confident in their naïve conceptions [33]. In cognitive science, this effect is long known as “proactive interference” [34] – meaning that if any information is only presented, even though it conflicts with a learner’s prior concept, it does not have a lasting impact in the long-term memory [35]! Previously learned concepts inhibit the learning of new concepts because prior concepts are often based on own experiences and conclusions.

Additionally, developing metacognitive competencies might be a key factor for academic success [36] because it helps to grasp the limitations and applicability of the one’s own mental models. However, to “attain this level of consciousness students have to experience a process of generalising the new concepts in a large number of different situations” [29, p. 275]. In this mindset, the emergence of naïve concepts appears as an artifact of a theory overload or as a lack of practice and reflection opportunities.

IV. OUTLOOK

Including further datasets, confidence levels, longitudinal and/or qualitative studies will most probably yield even more insights and allow for statistically more sophisticated methods. Also looking for language-specific effects might yield insights in the differences in learning culture. The item wording could profit from moderate modifications: Rephrasing answer scale option 2 from “some agreeable facts” to “mostly disagree” could result in a lower cognitive load, more equidistant distractors and hence in less ambiguous answer patterns. Rationale #2 (skipping stone) should be differentiated more and be sensitive to whether students think particle paths will cross and if there is “vacuum” behind the wing. Also, the qualifier “fast” may trigger different concepts and should be deleted in further studies. Rationale #5 (Coandă effect) needs to be accompanied by a rationale specifically asking for the viscosity and/or friction of air. Rationale #6 (circulatory flow) needs a clarifying picture and should be contrasted with wingtip vortices. Maybe by the help of another rationale specifically asking for the contribution of wingtip vortices to lift. Further implications for learning and teaching can be found at Potvin [38].

ACKNOWLEDGMENTS

Florian Genz would like to thank Ben Archibeque, Caroline Böning, Zoë Bohlmann, André Bresges, Cornelia Genz, Kristina Vitek, Rainer Zimmermann, and the PEER project (Professional Development for Emerging Education Researchers) for valuable feedback. The Future Strategy for Teacher Education (ZuS) is part of the “Qualitätsoffensive Lehrerbildung”, a joint initiative of the Federal Government and the Länder which aims to improve the quality of teacher training. The program is funded by the Federal Ministry of Education and Research. The authors are responsible for the content of this publication.


