

# Questions and answers about the Nature case:

Which “new” pieces of evidence were brought forth in Spring ‘21 by the Three Physicists?	1
How did the National Committee for Scientific Misconduct rule in Nov. ‘21?	3
What did the AU Research Practice Committee conclude in May ‘20	3
Did Dries Sels in 2017 ask for the erroneous algorithmic code?	4
Should we have known that Dries Sels in fact meant the algorithmic code as the AU Committee has ruled in 2020?	5
What is the difference between the game code that Dries Sels mentioned in 2017 and the erroneous algorithmic code?	5
If the May ‘20 AU research practice ruling is factually wrong, why did you not object at the time?	6
Have we ignored legitimate scientific criticism that our research is “untenable” over a period of several years?	7
What did scientific peers think of Dries Sels’ “untenability” criticism of our work?	7
If there was a sign error in the optimizer code causing it to worsen the solutions rather than improve them, how could we optimize the player solutions so well with that very same algorithm?	8
Is it correct that three critics who have raised legitimate scientific criticism have been persecuted by the AU leadership and external foundations?	9

If any important questions are left unanswered, please email [sherson@mgmt.au.dk](mailto:sherson@mgmt.au.dk)

## Which “new” pieces of evidence were brought forth in Spring ‘21 by the Three Physicists?

Following the harsh and one-sided media coverage of the case by the Danish newspaper Berlingske in Dec. ‘20 and Jan. ‘21 and the ensuing popular outrage, the Dean of the Natural Science Faculty at AU made a public announcement that any further allegations had to be brought forth by Feb. ‘21 in order to be considered. The Three colleagues at my Physics department took the occasion to formulate two “new” questions to be investigated:

- Have I known about the error since summer 2017 and concealed this knowledge from the public and scientific colleagues?
- Were funding applications made in that time that were based on Nature paper results to be valid?

Since I never actually touched the relevant codes, of course I could also not be aware of a mistake in the code alone. Therefore, the allegations are really that a dozen of students and programmers who have been in touch with the code over the past 4 years and I should all have

a shared conspiracy to conceal that knowledge to the public even though the vast majority of them are now in industry jobs and never have had any personal motivation to keep a secret like this concealed. So, if that conspiracy theory were in fact true, it would be one of the most spectacular cases of systematic and coordinated fraud in the Danish history of science. An allegation of this magnitude definitely deserves to be taken seriously but at the same time it should also be considered that these kinds of allegations are not just career damaging to the involved researchers but also often career ending. So what were the compelling new pieces of evidence presented by The Three?

First, as mentioned above, The Three had, in the summer of 2020, after the ruling of the RPC, found evidence from a public code repository and master's theses from my group that we had implementations of the algorithm in C++ without the numerical error as early as end of 2016 and they conclude that this documents that we must have found the error then, corrected it and concealed that knowledge in the coming years. However, in truth this was the very same algorithmic development that I openly mentioned to Dries Sels in the summer of 2017, thoroughly explained in the report to the RPC in the spring of 2020 and [explained satisfactorily in detail to the Head of Department](#) and sent to The Three in the autumn of 2020 without them caring to comment on it.

But that was of course not the only "new" evidence. The next argument was a single sentence extracted by Berlingske from an email that I sent Flemming Besenbacher on the 24th of January 2020 to inform him that we found numerical error in the simulation code. This is sensible, since he is chairman of the Carlsberg Foundation, which is currently our main funder. I wrote "There is a small but very important numerical error in our original code... We have long known that other algorithms can indeed solve the problem. We have therefore throughout the last 3-4 years built that part of the project [Quantum Moves] around a much more humble interpretation of our results". Is this statement proof that I have through the process redefined what the Nature paper proved as the critics claim and should have retracted the paper already in 2017? No. There is a dramatic difference between what a paper demonstrates and how one can subjectively interpret the generality of these results. As we started to develop more standard algorithms that could handle the problem, my own private interpretation of the results changed from believing that standard algorithms *would* fail on the problem to that they *could*. Since we had been consciously cautious about reporting only our observations relative to the single employed algorithm and refrained from making bold generalizations in the paper, this meant that *at the time* there was no reason to bring the particular reported results into question.

So, the summary is that The Three again only have loose circumstantial evidence that can however be stitched together to provide a compelling "it sounds too incredible to be true" case if one is sufficiently selective in the picking out of the facts of the story. Once again, I wrote a report trying to set things straight and sent it to the AU Committee as an initial response to be considered together with the formal complaint put forth by The Three as the committee decided whether to take up the case again or not.

I would later find out that the official allegations contained one more "smoking gun piece of evidence" that had been added *after* I had been allowed to read and comment on their first draft. The accusers had *added* a single sentence from an email I'd written in 2017 to Dries Sels: "except for an annoying numerical mistake, we would have put this work on the arXiv...". Rather than being a "smoking gun" piece of evidence of their conspiracy theory, in reality it was taken

completely out of context! The reference was to a mistake that had nothing to do with the Quantum Moves investigation. Rather, this was about a small numerical error made by a German research team in one of their papers. As part of our algorithmic development we were replicating the results from their paper as a test of our new algorithms. We wanted to ensure that we were using exactly the same parameters as they'd used in the original work. Over the course of an in-depth, collegial dialogue with the involved professor, they verified that they had, in fact, accidentally reported the wrong numerical value in the published paper. We corrected for it and continued with our work, albeit with many months delay, meaning that our work was not ready for submission until February 2018 rather than summer 2017 when the discussion with Sels was ongoing and the quote was made.

Feeling disheartened by the fact that, once again, these colleagues had not even asked me to clarify directly, assuming quite deliberately and dishonestly the worst instead. I also wondered: why wasn't this included in the first version to which I had written my response? And as they revised their complaint, why was I never given the opportunity to respond to this new accusation?

## How did the National Committee for Scientific Misconduct rule in Nov. '21?

I have previously explained how the harsh allegations of my three physics colleagues in early '21 contained no new pieces of evidence. The case was with the AU committee for a few months after which they decided to send it to the National Committee for Scientific Misconduct. In Nov. '21 I finally received their ruling. To my great relief the committee unanimously rejected all allegations of scientific misconduct. More explicitly, they ruled that the fact that there was criticism of the scientific article in 2017 does not in any way imply malintent or fraud in our subsequent behavior.

I drew a huge sigh of relief and happily informed my employees and close colleagues that the harrowing ordeal now finally seemed to have come to a conclusion and we could start to move on. Alas, again the relief was only short lived because short before christmas I was informed that my university would now try to see if the latest round of dirt-digging could potentially give rise to new accounts of questionable research practice on top of the old convictions. In every previous investigation I have welcomed the occasion to tell my side of the story and get the truth out, but I sadly no longer harbor any optimism for the fairness of the process. The latest investigations are supposedly set to end in January '22, but I have no illusions that the treatment will be any more fair than the rest and neither that they will signal the end of my personal and professional nightmare for which there is unfortunately no end in sight.

## What did the AU Research Practice Committee conclude in May '20

Despite a long list of very harsh allegations from three critics at our own physics department including that the error was created on purpose before the time of publication and having knowingly *"attempted to conceal the errors by leading attention away from sensitive questions"* the Committee found no indication of scientific misconduct such as *"purposefully having*

*attempted to withhold important research data or other pieces of information in order to hide or conceal their mistakes*". The Committee also did not find "any reason to assume that JS or KM in any phase of the research work - also not after the criticism had started - have behaved dishonestly" or "generally rejecting to engage in discussions of scientific matters". Generally the Committee finds that the matter "was characterized by a scientific communication gone bad and that has created personal mistrust among the involved parties". Concretely, the Committee issued a ruling on two issues of potential questionable research practice:

- Delays in handing over the numerical codes. They did, however, conclude that I should have delivered the simulation code to Grønlund sooner than early december '19 following an initial request in July '19, an offense the committee labeled 'questionable scientific practice'. Even though the documentation and validation of the codes did take a significant effort to prepare and I at the time thought I did my best to clarify the scientific dispute by working on a short public response (requested by my Head of Department) and a more thorough validation study, I acknowledge that the handover of the code should have taken precedence. With a concentrated effort, the codes could have been handed over in September or October and I take full responsibility for the delay until Dec 1st and my error in judgment. However, much to my dismay, the Practice Committee also issued a judgment that caught me completely by surprise. In their ruling, while the Practice Committee acknowledged Sels had asked me for the game code, not the simulation code, nevertheless they ruled that I should have somehow intuited that Sels might also have wanted the latter and delivered it although he had never implicitly or explicitly requested it! See appropriate Q&A entry for my reasons why I consider this ruling completely unfair. Briefly, Sels himself wrote in a public manuscript in 2018 that he found no reason to suspect any error in our codes!
- Passive behavior wrt self-control of codes: the Committee found that our behavior "might be considered insufficient effort, but that it was not likely that it in itself with sufficient certainty can be considered a breach of responsible research practice" but that it could be added as a supplemental element in the complete evaluation. Now, it is important to understand that mistakes in numerical codes are not a matter of finding the one mistake or not. In generating codes there are always a multitude of different numerical errors that are identified until sufficient trust in the code is established to report the results. We have throughout the entire matter openly presented our code checks and further development efforts and although our opponents cherish in explaining how easy the mistake was to find, the fact remains that it is much easier to make such arguments in hindsight that we did not possess as we developed the code. It would therefore be a strange counterfactual process to convict us for behavior in 2016 based on knowledge in 2020.

## Did Dries Sels in 2017 ask for the erroneous algorithmic code?

No! He asked a concrete question about the game window of the players. In that connection he asked for an antiquated game code. Since we knew that this code could not answer his concrete question, I immediately answered his specific question and he has never in any of our further emails over the coming 4 years repeated a code request. I therefore was fully convinced

that I had answered all of his requests until I to my shock saw the accusation presented by our Head of Department in the letter to the AU Committee for Responsible research that we should have refused to hand over a code since 2017. For the difference between the mentioned game code and the erroneous algorithmic code, see the appropriate Question.

## Should we have known that Dries Sels in fact meant the algorithmic code as the AU Committee has ruled in 2020?

No! We were in 2020 convicted for questionable research practice among others on the grounds that the Committee ruled that, whereas Dries Sels did not ask for the erroneous codes, we should have taken his persistent criticism of our work as an indication that he really wanted the algorithmic codes. We disagree enormously with this part of the ruling for three reasons:

- We responded immediately to his questions in 2017 and never heard a repeated mention of code request.
- Independent experts already in 2017 disagreed with the criticism of Dries Sels and called his presentation both unsurprising and unfair. Thus, the interpretation of the criticism at the time was not as obvious as our current critics in their clear hindsight want to make it seem like. Neither experts, nor very importantly Dries Sels (see below) at the time suspected the validity of the codes and it is therefore an unfair counterfactual judicial verdict to demand that we should have been able to foretell the error, when no one else could.
- Most importantly, Dries Sels did not at any point bring the validity of the algorithms into question! Explicitly, in his [2018 paper](#), he claimed that the superior performance of his algorithm over ours stemmed from a numerical trick he had developed that was not implemented in our algorithmic codes. "Note that the method is relatively efficient, since we can compute the local change in the fidelity (step 3) for all tweezer positions without having to recompute the entire time evolution (step 2), in stark contrast to the KASS method used in Ref. [8]. Since the latter works in Fourier space, one must recompute the entire time evolution for every small change to the parameters." In summary, in complete agreement with our assessment at the time, Dries Sels in 2018 believed that our (KASS) algorithm was correctly implemented but lacked the numerical sophistication of his own algorithm.

## What is the difference between the game code that Dries Sels mentioned in 2017 and the erroneous algorithmic code?

In short, there are two different kinds of computer codes in the case. First, there is the code that is used to define how the game should look and respond: the game codes. These control how each level should look like and what should happen with the sloshing liquid when players move the mouse. This code is without errors and it is a very early version of this game code that Dries Sels referred to in 2017. This fact has been acknowledged by both the AU Committee and in the [AU official timeline](#) (spilkode versus "koder" 12/7-19), although the opposite unfortunately was claimed in the original case description put

forth to the Committee. So, the game code is the one the player interact with directly and the one we used to [gather data](#). All plots of this [raw data](#) depicted in the original Nature paper is still valid. After having collected the data we needed to [analyze](#) it. This was done using the algorithmic codes. These codes were used to both to iteratively improve the players' raw data as well as a set of randomly generated raw data to minimize the sloshing in the target area. The comparison of the end result of this algorithmic fine-tuning between player and random inputs is the main result in our original Nature paper, where we demonstrate that the player input is a better starting guess than the randomly generated ones. This conclusion has been validated in our recent work, but unfortunately the algorithmic code contained errors, which meant that the quantitative results no longer hold true and the paper accordingly retracted.

## If the May '20 AU research practice ruling is factually wrong, why did you not object at the time?

As explained elsewhere, the committee found that we should have handed over codes in 2017 that Dries Sels never asked for and that we can document in writing that he did not suspect to be erroneous at the time. Why then did we not challenge the ruling in May '20? I was presented the committee ruling by the Dean of the Natural Science Faculty and head of department of Physics. I informed them that i disagreed strongly with the wording of parts of the committee ruling, but in response the AU leadership guaranteed me that all involved parties would be informed that this was to be considered the end of the matter. After several years of intense pressure, many sleepless nights, stressful inter-departmental meetings, and countless clarifications, it felt as if the dark clouds were finally starting to lift. It was time to move on. We did not feel that we had been judged fairly but at least we would now get peace. Or so I thought...

It quickly turned out that The Three from Physics were not happy to let the matter rest. They kept digging and spreading rumors in both scientific and funding communities. Their increasingly preposterous and thinly documented allegations have successfully kept the case alive for another year and a half, so far. Their allegations of scientific misconduct have been unanimously rejected by the national committee in Nov. '21. Before that, their allegations had been tested internally at the department but apparently not in a satisfactory way because in the fall '20 someone anonymously contacted the Danish newspaper, Berlingske, with exactly the same loosely founded allegations. Then came the time in which i bitterly regretted not having challenged the AU committee ruling: Berlingske used as their main angle that we had for years arrogantly rejected rightful scientific criticism and refused to hand over codes since 2017. All of this is completely wrong but the newspaper could use the formulation for the AU committee as witness in their case. To me, this clearly illustrates that the current system is not equipped to handle cases in which systematic bullying is mixed in with ongoing scientific disagreements because the committee refuses to engage in either of these topics.

## Have we ignored legitimate scientific criticism that our research is “untenable” over a period of several years?

No. We have throughout the years received advice from leading experts in the field and independent advisors and their conclusion has been: there is nothing surprising in Dries Sels study from 2017. We were advised not to engage in a round of public mud slinging and that instead the most sustainable response would be to conduct a new study involving more games and comparison to more algorithms. This study is now [published](#) and confirms the main point of the original work, and therefore in our opinion constitutes a full scientific response to the criticism raised.

## What did scientific peers think of Dries Sels’ “untenability” criticism of our work?

In short, they found it unfair and manipulative. During the summer of ‘17 Dries submitted his work to the prestigious journal *Nature* in which our original work had been published. Here, the refereeing process is extremely thorough. After an initial editorial filtering the manuscript was sent to three anonymous, expert referees and to me for comments. The expert peer reviewers were very critical and commented on his manuscript that it was “sloppily written”, “not that surprising” and that it contained “several strong claims, regarding comparisons to prior work, that appear unfair” and unanimously rejected his manuscript from *Nature*. One referee pointed out that the only fair way to determine a dispute about the relative importance of human and algorithmic contributions to quantum optimization was to test the algorithms at the same time or before and definitely not after first results from the players had been reported and the good solutions were publicly known. As we shall later discuss, we took this expert advice on the best way to settle any disputes and therefore began developing Quantum Moves 2 with a broader suite of challenges and algorithmic comparisons. In private communication with me, Dries Sels admitted that his tone had been intentionally confrontational (“I’m well aware of the confrontational tone of the draft, and apologies for this.”, email 14/9□17) and he promised to give a more balanced account of all of our results in future submissions of his work (“It’s interesting that the HILO is comparable in speed to the stochastic optimization. I will see if I can mention this somewhere in the revised version” email 10/9□17). After the summer ‘17 we did not hear more from him and big was our surprise in April 2018 when we found out that despite his promises had submitted a largely unchanged manuscript to a much more specialized journal, *Physical Review A*, and apparently gotten it passed the single peer reviewer. His misguided methodology (in that his results excluded all but one of our algorithms) and argumentation were now out in the open as a peer reviewed published [heavily criticized by our faculty’s research practice advisor](#), and in 2019, a committee of quantum control experts evaluating the PhD thesis of the first author of the *Nature* paper noted a strong criticism of Sels’ work: “it contains a very mature discussion of criticism brought forward by someone else [Sels], whose communication standards seem questionable at best”.

If there was a sign error in the optimizer code causing it to worsen the solutions rather than improve them, how could we optimize the player solutions so well with that very same algorithm?

To us, the fact that players alone outperformed the purely numerical algorithm (red curve) was interesting but of no practical importance, since much better results were obtained by optimizing the player solutions, as we demonstrated with the blue and yellow curves in the Nature article. It was the comparison of the blue and yellow curve to the red curve which was the most important result. It was believed that the same optimization algorithm was applied to both the randomly generated solutions (red curve) and to players generated solutions (blue and yellow curves). Even if it later seemed that the optimization algorithm used by us was not the most efficient - since our follow up work and the work of others were able to obtain better results when optimizing random solutions - it remained remarkable that the old algorithm was showing great difference in performance when applied to the two types of solutions, implying that the human input can make a great difference in the final outcome. This was an important finding, because it could plausibly occur in other quantum processes, even if one had the best algorithm available. As it later turned out, the algorithms applied on the two types of solutions were not equivalent! **The sign error** was present in both cases, but it **only affected optimization of the random solutions** (red curve). In the case of player solutions, the optimization of the tweezer depth (where the sign error was made) was inadvertently disabled by changing a setting of a so-called step size parameter to a very small value ( $10^{-4}$ ), implying that the tweezer depth was not optimized at all. In the case of random solutions, the step size parameter was set to 0.1, which is also quite small, given the tweezer depth is of the order of 100. So even though the tweezer was made shallower instead of deeper, it happened only very, very slowly and the algorithm could still compensate by optimizing the tweezer position. It was in fact still so efficient, that it could optimize several long duration solutions to a perfect fidelity. We can now only speculate how this error occurred. It was definitely **not** a conscious choice to effectively disable the optimisation of the tweezer depth for the player seeds, since we explicitly stated in the article that both the tweezer depth and tweezer positions were optimized. In simpler terms we could state it as follows: "There was a sign error, but it had a detrimental effect only on the optimisation of the randomly generated solutions. It did not affect performance of the algorithm when applied to player solutions because of a different parameter setting which counteracted the sign error. The algorithm therefore worked much better on the player solution than on the random solutions. The fact that the algorithm could greatly improve the human solutions as well as the random solutions with long durations, led the researchers to believe that it was working properly and they attributed the difference in performance between the pair of yellow and blue and the red curve to the quality of the human solutions."



# Is it correct that three critics who have raised legitimate scientific criticism have been persecuted by the AU leadership and external foundations?

## Examples of statements

- “Olav W. Bertelsen hopes through discussions in the main leadership-employee forum to get assurance that rector and the remaining leadership has the back of the employees - also when the raise scientific criticism” [Omnibus 19/1](#)
- “How to balance the need for external funding against the need to protect the right to open scientific criticism and their employees.” [Nina Koefoed 29/1 Berlingske](#)

The case on the retraction of our Nature article has been used as an occasion to discuss freedom of research, the influence of external foundations on university policy, and correct handling of scientific criticism and errors in research. All of these topics are relevant and pertinent in the present day research environment. It is, however, unfortunate that these discussions are based on the one-sided hero-villain presentation of the matter presented by Berlingske when the truth is much more complex.

In order to better understand the matter, it is important to distinguish between two groups of critics in the present case: i) two researchers who have chosen to engage in a scientific debate by writing scientific articles on the topic and ii) three researchers at Physics, who have refused to engage in the scientific parts of the matter and only argued that it should be handled differently. The Three from Physics have have neither engaged in any form of scientific debate or requested any scientific documentation from me. They have told me that they have not even read the Nature paper and at the only officially arranged meeting between they explicitly said they “did not have any opinion about the scientific part of the matter” (official summary of the meeting 19/5-19). When the Dean and head of Carlsberg foundation called The Three unpleasant names it was therefore in no way a suppression of the right to scientific criticism but rather a frustrated response to the apparent collegial bullying that was being conducted by The Three. These statements indeed fell in the time just after the first ruling of the AU committee in may '21 after which the Dean urged all involved parties to bury the hatchet and move on. Something that The Three clearly refused to do.

We had (and may still have) a scientific disagreement with the two scientific critics. This is completely fair. We have responded to their criticism with a published paper and the two may choose to publish a response to this (which they have not don in the past two years since we put our response on the public arxiv). With The Three, the dispute has revolved around how we should respond to the two critics. The Three wanted a rapid official response to the criticism and we have chosen to follow the advice of experts in the field to “not engage in public mudwrestling” and instead conduct a more thorough scientific study. We have now published this work, so we have indeed followed the demands of The Three for a public response but probably not in a way that they had wanted. This difference in approaches has nothing to do with questionable research practice and the criticism of the bullying of The Three should not be mistaken for suppression of justified scientific criticism. On the contrary, in my view, this matter is much less about the right to *openly criticize* and much more about the obligation to *listen* when you have chosen to engage in such a dispute.