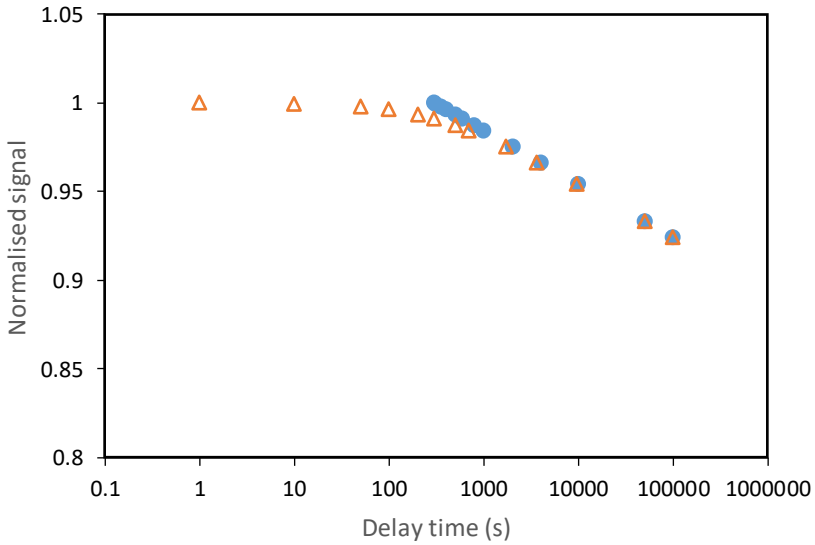
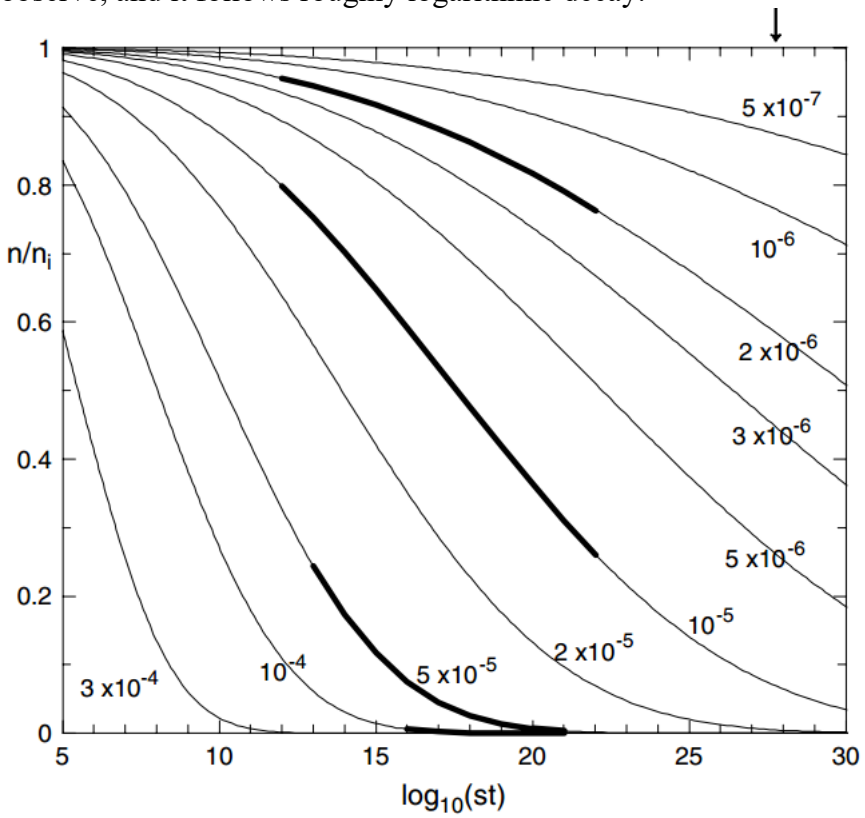
This paper reported some experimental observations about anomalous fading using two polymineral samples. The authors claimed that they have found different observations from what the majority of other people did, and provided comprehensive interpretation on their observations, and suggested a different way of doing fading tests. Although I understand that the authors spent a large amount of efforts in presenting the data and providing interpretations, I feel frustrated to try to understand and follow what exactly the authors have done. The manuscript was written and organised in a complicated way, which prevents a good flow of the contents. If I understand the authors correctly, I am afraid of that their observations are built on a wrong way of data presentation, so most of their explanations and implications for dating are not supported.

Major issues:

1) One of the major issues is that the authors presented their fading data in a wrong way. They artificially set the delay time of the first measurement (prompt measurement) to zero. This will artificially change the shape of the fading curve (and fading rate). Let's take a simple example. Suppose that we have a sample with a fading rate of 3 %/decade, and we conducted a series of delayed SAR measurements with irradiation (400 s). The IR-readout or delay time after the mid-point of irradiation is 300, 310, 350, 400, 500, 600, 800, 1000, 2000, 4000, 10000, 50000, 100000 s, respectively, which corresponding to about maximum of 2.5 decades of time (similar to most of fading tests conducted in previous studies). Supposing that its fading follows a logarithmic decay, we can theoretically calculate their decay curve (see the blue dots in the following figure). However, if we artificially set the initial decay to 1 s (btw, I don't know how could the authors show the first data point in a log scale, if the first time is 0 s, do you just ignore the first data in the figures?), what happens is that the decay curve (orange triangles) is flatten out at short delay times (<300 s), which is exactly what the authors observed in their data sets. This is actually an artificial result.



2) If my above comment is built on misunderstanding of the authors way of presenting data, and supposing that the author's presentation is correct (e.g., if we can indeed measure the fading curve down to a few seconds after irradiation, e.g., assuming that we can apply a short pulse of large dose), then do we expect such a pattern for real sample? The answer is yes, and fading curve (signal loss) should not follow a logarithmic decay function. The theoretical decay of signal should actually follow a power-law as proposed by Huntley (2006) in their Figure 2 (see below). If we compare the theoretical curves with the patterns shown in this study, they are remarkably similar, with a flat decay followed by a sharp decay and then flat decay again. The reason why we always assume a logarithmic decay is that, in practical, one cannot measure the signal at the beginning of the decay (time too short) and the end of decay (storage time too long). The middle part of the power-law decay is what we can practically observe, and it follows roughly logarithmic decay.



3) Should we apply preheat immediately after irradiation or after delay for fading test? My answer is "both are incorrect but a prompt preheat is definitely better". We already know that, fading rate is strongly dependent on the charge concentration and distribution in the lattice (e.g., Huntley 2006; Li and Li, 2008). A prompt preheat is the best way to mimic the electron-hole distribution for natural samples, whose thermally unstable (and athermally unstable) electron-hole traps remain empty. Of course, one would expect even larger fading in natural process, because the electrons can undergo multiple filling and escaping process (hence more fading), whereas they only have one-off filling and escaping in the lab fading test (because irradiation time is much shorter than storage time, and the signal loss can only be monitored after irradiation). So, the authors' claim of "preheating prior to IRSL-readout of the natural dose occurs after a long storage in nature" is wrong. Natural process is a long-term 'preheat' at low temperature. Kinetically, it has a similar effect of short-term preheat at

high temperature. That's why we need to apply preheat in the laboratory to mimic such process. Fading test, therefore, should not be conducted by keeping these unstable traps filled (e.g., using delayed preheat).

4) The authors discussed the fading test results without considering the fact that D_e measurements usually have the same parameters used for fading tests, such as stimulation temperature, and the magnitude of delay time between irradiation and IR measurements. So any effect of changing in measurement conditions for fading test, may also influence the De estimates. For example, let's take two scenarios. The first scenario is that both fading test and De were measured by using room temperature IR stimulation, and the 2nd scenario is that they were measured at 60 degree C. If one get a higher g-value when room-temperature IR readout is applied, then the corresponding De would probably lower too. That means, one can probably get the same fading-corrected results no matter what readout temperature they use and what fading rates they get. Of course, whether fading correction is reliable is another topic, as we need a model to describe the fading process, which is poorly known unfortunately.
   This issue also applies to the other experiments described by the authors, e.g., heat input, pause position, liftup temperatures, etc. Ultimately, the authors should test whether these conditions also influence De results or not. If they do, will you get the same results after fading correction? If you do get the same results, then that suggest no matter what fading rates you get, they are reliable.

Minor issues:

1) Section 1.6. The authors used some example of pIRIR results to support that fading measurement cannot provide insight to the stability of pIRIR. However, they ignored the fact that there are many other effects contributing to whether a pIRIR age is consistent with 'expected age' or not, such as residual correction (see the recent paper by Rui et al. QG, 2020), initial sensitivity changes (e.g., Qin et al. QG; Zhang et al., QG). They also ignore that not all the pIRIR results are reliable (see Li et al., 2014 Geochronometria for a review). As the other reviewer suggested, if one wants to study the effect of fading, they need to do a factorial experiment, i.e., ensure that only one parameter (fading) can affect the results. But I doubt this is possible at this stage, given the complex processes involved in IRSL dating.

2) Section 2.2. I am confused about why the authors did not use the option "run one aliquot at a time". My understanding is that, if they chose not to use this option, then the timing between irradiation and IR readout would vary from aliquot to aliquot (if they input the aliquot positions in the same sequence for each row). Then this would cause problem in SAR procedure, as different aliquots would have different delay time between the test dose irradiation and test dose signal measurements, which can result in problems with sensitivity corrections using test dose signals (as the test dose signal will fade differently for different aliquots).

3) The authors conducted many different version of fading tests with different parameters. I found it is extremely frustrating to try to understand what exactly they have done by reading the texts (section 3). I would suggest that the authors

provide a flow chart or sequence table (such as those commonly used to describe SAR procedures) for each of the fading tests, which will greatly aid the readers to understand details about their experimental procedures.

4) Is NRM the same as test dose? If yes, please use test dose, to keep consistent with the term used commonly by others.

5) The authors' results are based on polymineral, which contains all kinds of feldspars. What happens if Na-feldspar, Ca-feldspar and K-feldspar fade differently and respond to different stimulation conditions differently? For example, if some of them has an extremely high fading rate (say, ~20%, then the overall effect is that the fading decay will be flatten out in the later part of delay as the fast-fading component become substantially small than stable components.

6) Figure 1. I have no clue what the data points or model represents. What do different colours and symbols represent in 1, 2 and 3?