1. I would caution against the final sentence of the abstract, which says "these results show that high NA lenses are not an essential requirement for optical tweezer experiments." Certainly this statement needs further qualification. At face value, I think it is widely known that high NA is not strictly necessary for single-atom control and detection. However, for most Rydberg-based physics goals - particularly outside of the blockade regime - a high degree of atomic localization is essential, and this requires tight optical traps generated with high NA objectives. For assembled Hubbard systems, the requirement on high NA is thought to be even higher since tunneling must be controlled. My point is only that it depends on the goals of the experiment.

We have removed the final sentence of the abstract.

2. In the first sentence of the third paragraph in Section 4, the authors say "very deep tweezers". Please add a number here, even if it is rough. The authors use depths in uK units later in that paragraph, so I would like a similar estimate to quantify "very deep".

Here we consider a tweezer to be deep when the differential AC Stark shift is much greater than the linewidth of the imaging transition, rather than using an arbitrary temperature scale. The corresponding definition has been added to the paper.

3. The discussion in that same paragraph on subsidiary intensity maxima is a bit surprising to me. Can you quantify the relative depth of these additional maxima? Is that estimate consistent with the alleged Airy disk pattern induced by the circular aperture of the lens. What about the size profile? The authors suggest they have some spatial resolution of this effect. More quantitative detail would be helpful here since this is perhaps a new observation for Sr.

Modelling the system using Zemax we expect the subsidiary maxima to be approximately 1 % of the peak height, which at the trap depth of 30μK gives subsidiary maxima of the order of the MOT temperature. We have added these numbers to the paper. The evidence of this trapping is discussed in further details in the thesis of Ryan Hanley [29], which we have added another citation to later in the paragraph to ensure that proper attention is drawn to it.

In a ballistic expansion, we are able to image two expanding clouds due to the different expansion rate for each temperature. We have added a sentence to the paper.

4. A few paragraph later, starting with "A typical release and recapture signal...": what is the trap depth corresponding to that temperature measurement and estimate? Would it be appropriate to estimate radial n (motional quanta)?

The corresponding trap depth 520 μK, as stated in the caption of figure 3a. This value has been added to the main text.

The radial $\langle n \rangle = 8.6$, though we feel temperature is the more appropriate value for use in the body of the paper.

5. In Figure 4, what is 'N' as the vertical axis of the inset plots? Presumably This is related to atom survival, as described in the text. Perhaps the axis should be normalized to unity?

N is number of atoms in the trap, based upon the detected number of counts. The caption of Figure 4 has been updated make this clear. The data were taken in the multi-atom regime, so atom number is more appropriate than a survival probability, which could imply there is only one atom present. We considered normalising the axis, however since the maximum value of N actually seems to occur at different times (compare the inserts) the choice of an appropriate N_0 is not straightforward.

6. In Figure 5a, I would suggest adding the relative contribution of the four sections. Presumably this is consistent with the Nbar=1.2 stated in table 1.

We added the relative contribution of the sections to the figure as suggested. However the addition cluttered the figure and detracted from the histogram, so we do not include the relative contributions in the final version. It is consistent with Nbar = 1.2.

7. In the discussion of pulsed detection where blue scattering is pulsed and red Sisyphus cooling is always on: My understanding of the initial observations of this "repulsive" Sisyphus mechanism is that it can repel the atom both below and above a critical energy. Is it possible that the blue scattering quickly heats the atom above the critical energy with some probability, after which it is lost since the "cooling" only hurts in that case? Further, if the atom is not near the critical energy, the "cooling" largely does nothing. Is it really necessary to use the "cooling" during the time when the blue pulse is off?

This mechanism where the atom is not captured by the cooling is almost certainly why, as we state in the text, the cooling is not optimal, and we still see some heating off resonance. Varying the Sisyphus detuning during the imaging sequence may help.

The cooling does nothing during the blue pulses, as they are so short that almost no red photons are scattered. The cooling only works during the much longer dark periods. Indeed we observe that without red light during these periods there is no cooling.

8. Finally, I believe the reader would appreciate a stronger and more clear final outlook. What exactly are the intended goals of this experiment, and what further directions would benefit most from its unique features? The authors mentioned precision optical metrology in their response letter, but no such statement is explicitly made in the text, although it is perhaps implied in the last sentence.

Other possibilities for future work include, but are not limited to, precision measurement of Rydberg states, and studying spin and charge transport. These have been added to the outlook with appropriate references.