For research article

Response to Reviewer 3 Comments

1. Summary

Thank you very much for taking the time to review this manuscript. We have carefully addressed each of the issues you raised and implemented revisions based on your valuable suggestions. Detailed responses to your comments are provided below, with corresponding revisions in the re-submitted files.

3. Point-by-point response to Comments and Suggestions for Authors

Comments 1:

There were so many factors in the experimental design: breed, number of pigs, the number of pens to be mixed, and which day after mixing of a pig to be introduced. But there were only 6 groups in the experiment, and each treatment had only one repetition (pen). It is suggested that the authors reduce the factors analyzed in the experimental design.

The design of mixing a pig into the population every three days is also problematic, because the result of the later time point is influenced by the previous time point. Then the results of each time point cannot be simply compared. The authors should only compare the differences between groups.

The shortcomings of these experimental designs should be discussed in the discussion.

Response 1: Thank you for pointing these out. We agree with these comments. Therefore, we reduced the number of analysis factors in the experimental design and made revisions accordingly. Following these changes, we conducted a reanalysis of the data and updated the results section accordingly. All revised statistical findings and other pertinent modifications resulting from this reanalysis are now highlighted in red throughout the revised manuscript. Furthermore, we have included a thorough discussion on the limitations of our experimental design to provide context and emphasize the scientific implications of our study.

- **The modifications to the experimental design are as follows:**

During the pilot test, we also found that the number of resident pigs involved in aggression in the herd varied with T and that the aggression intensity of the resident pigs varied with the number of pigs from different pens. Accordingly, we hypothesized that AI would be related to T and the structural composition of the herd (including size and the number of pens).

Subsequently, we designed two experiments: Experiment 1 and Experiment 2. Fig. 2 presents a schematic representation of the experimental design.

In Experiment 1, we investigated the impact of size and establishment time of the herd on the aggression intensity towards newly integrated unfamiliar pigs. We created 4 groups, including two control groups and two treatment groups. Control 1 (CTRL1) consisted of 9 pigs from two pens, and Treatment 1 (TERAT1) consisted of 12 pigs from two pens. Control 2 (CTRL2) consisted of 9 pigs from two pens, and Treatment 2 (TERAT2) consisted of 12 pigs from two pens. Following, T was set with 7 gradients for the groups above, 3 d for each

gradient interval, T = 3 for the initial value, and T = 21 for the cutoff. At T = 3, 6, 9, 12, 15, 18, and 21, unfamiliar pigs were integrated into each herd. To ensure accurate measurement of aggressive behavior and minimize interference in determining the effects on pig group dynamics, we opted to add only one pig into each group at each time point.

In Experiment 2, we investigated the impact of the number of pens composing the herds and herd establishment time on the aggression intensity towards newly integrated unfamiliar pigs. There were 4 groups (2 of the groups were from Experiment 1) in Experiment 2, consisting of 2 control groups and 2 treatment groups. The control groups (CTRL1 and CTRL2) each consisted of 9 pigs from two pens, and the treatment groups (TERAT3 and TERAT4) each consisted of 9 pigs from more than two pens. Following, T was set with 7 gradients for the groups above, 3 d for each gradient interval, T = 3 for the initial value, and T = 21 for the cutoff. At T = 3, 6, 9, 12, 15, 18, and 21, unfamiliar pigs were integrated into each herd. Photos of the six groups of pigs are shown in Supplementary Material Figure S1.

Experiment 1 and experiment 2 were performed for 42 integrated pigs and 6 pig herds in total. For each series of experiments, 7 pigs were integrated into the same herd at 7 different times. Each integrated pig was unfamiliar with the resident pigs, while its breed was consistent with that of the resident pigs. Familiarity at this stage was defined as sharing a previous pen with no regard to whether they were littermates. For each integration, a professional breeder and a professional veterinarian were involved in addition to the experimenters conducting the study. Based on actual production experience, breeders integrated pigs into new herds without violating the guidelines for the welfare of commodity pigs. In accordance with the biosafety requirements of pig farms, veterinarians disinfected and treated pigs with moderate and worse skin lesions.

This change can be found in the revised manuscript on lines 171–215.

- **The main modifications to statistical analysis are as follows:**

We performed normality assessments on the six datasets under scrutiny. The outcomes derived from both the Shapiro–Wilk and Jarque–Bera tests demonstrate that all six datasets display characteristics consistent with a normal distribution. Specifically, the p-values yielded by the Jarque–Bera test exceed 0.05, indicating the absence of significant deviations from normality within these datasets. These results affirm the premise that the data utilized in our investigation conform to a normal distribution. For herd size and the number of pens from which groups were created, we conducted paired t-tests to analyze the intergroup differences in pig group structure on the intensity of aggression. For the time to establish the group, we performed a nonlinear regression analysis to assess the relationship between the time to establish the group and the intensity of aggression. Considering the impact of the within-group factor T, we first used the four pig groups created in the group size experiment as sample IDs, treating the time to establish the group as a within-group item, and CTRL1, TREAT1, CTRL2 and TREAT2 as between-group items. We performed repeated measures ANOVA with the intensity of aggression as the dependent variable to determine whether there was a differential effect in the interaction between within-group and between-group factors. Subsequently, we used the number of pens composing the pig groups as sample IDs, establishment time within each group as an intra-group factor, and CTRL1, CTRL2, TREAT3, and TREAT4 as inter-group factors to conduct intra-group effect analysis.

This change can be found in the revised manuscript on lines 379–397.

- **The main modifications to the result section 3.1 are as follows:**

First, we measured the AI of CTRL1, TREAT1, CTRL2 and TREAT2 under different T conditions within their groups. These AIs showed no significant differences between the different pig groups (Table S2). This indicates that the intensity of aggression towards unfamiliar pigs is not statistically significant between groups of size 9 and size 12 when unfamiliar pigs are integrated into the group.

Subsequently, we measured the AI of CTRL1, TREAT3, CTRL2 and TREAT4 under different T conditions within their groups. These AIs were different (Table S3). Among them, significant differences were found between CTRL1 and TREAT3 ($t = 8.000$, $p < 0.001$; Fig. 3a), between CTRL1 and TREAT4($t = 7.071$, $p < 0.001$; Fig. 3b), between CTRL2 and TREAT3 ($t = 6.000$, $p = 0.001$; Fig. 3c), and between CTRL2 and TREAT4, ($t = 4.382$, $p = 0.005$; Fig. 3d), with the mean AI in CTRL1 (3.57) and CTRL2 (3.29) being significantly higher than the mean AI in TREAT3 (2.43) and TREAT4 (2.14). This means that the intensity of aggression towards unfamiliar pigs is statistically significant between pig groups composed of 2, 4, and 6 pens when unfamiliar pigs are introduced into the group.

This change can be found in the revised manuscript on lines 405–418.

- **The main modifications to the result section 3.2 are as follows:**

To evaluate the impact of the within-group factor T on the intensity of aggression and to explore whether this impact is influenced by group size and the number of pens composing the group, we used a parabolic model based on the Levenberg–Marquardt method to perform regression analysis on the T and AI of six different pig groups.

This change can be found in the revised manuscript on lines 426–429.

- **The main modifications to the result section 3.3 are as follows:**

Considering the impact of the within-group factor T, we conducted repeated measures ANOVA on CTRL1, TREAT1, CTRL2 and TREAT2 to further analyze whether there is a differential effect in the interaction between within-group and between-group factors. The results showed that the interaction between CTRL1, TREAT1, CTRL2, TREAT2, and T was not significant ($p > 0.05$; Table S5). Subsequently, we conducted intra-group effect analysis on CTRL1, CTRL2, TREAT3, and TREAT4. The results showed that the interactions of CTRL1, CTRL2, TREAT3, and TREAT4, and T were significant ($p < 0.05$; Table S6). These results indicate that there was no difference between size. However, there was a coupling effect between the number of pens used to form herds and T. The interaction effect suggests that there is a significant difference in AI between groups formed with 2 pens and those with more than 2 pens when $T \leq 12$ ($p < 0.05$; Table S6). However, as T increased to $T > 12$, there was no significant effect of the number of pens used to form the herd on the AI.

This change can be found in the revised manuscript on lines 441–453.

- **The added content in discussion**

4.3 Experimental Design and Its Limitations

The selection of time points for the group establishment design (3 and 21 days) was based on considerations of the research application scenarios (the scenarios mentioned in the introduction). The design with a three-day interval was chosen to create a gradient of group establishment times, ranging from 3 to 21 days, to comprehensively assess the impact of group

establishment time on inter-group aggression. Additionally, the objective of this study was to explore the effects of group establishment time on aggressive behavior, thus requiring sufficient data support. Adding one pig every three days ensures an adequate sample size within the range of group establishment times for statistical analysis and reliable conclusions. Longer intervals may decrease the number of data points, reducing the effectiveness of statistical analysis and hindering an accurate assessment of the effects of group establishment time. Moreover, while shorter intervals may lead to pigs adding before the group is fully stabilized, this design allows for a more detailed observation of dynamic processes. This enables capturing changes and adjustments that may occur during the group establishment process, facilitating a more comprehensive under-standing of group dynamics and the mechanisms underlying aggressive behavior. In addition, this study faced extremely unique circumstances, leading to a series of experimental design challenges. Due to the lack of effective control measures, the study posed potential harm to newly integrated individuals, making it greatly restricted to conduct this research in commercial pig farms. Additionally, national legislation on animal welfare further restricted the scope of our study. These constraints severely limited our experimental design. Despite these challenges, we carefully considered the use of mixed models for more precise estimation of variance components and enhanced interpretation of experimental outcomes. However, during preliminary data analysis, significant collinearity issues among our variables were identified. Addressing these challenges was crucial as they could potentially affect the accurate estimation of fixed effects in mixed models, compromising their reliability and interpretability. Given these considerations, we opted for repeated measures ANOVA as a more suitable and robust analytical approach. Repeated measures ANOVA effectively handles within-group factors and provides a straightforward framework for interpreting the observed variance in our experimental design.

This change can be found in the revised manuscript on lines 541–570.