1. No, the ordinal regions in the normal mapped on the normal distribution (probit link) are not equally likely but differ strongly in their likelihood. When the stimulus gender is female, the likelihood for the estimates of the threshold coefficients are: P(1) = 0.005565486; P(2) = 0.1828717; P(3) = 0.331657; P(4) = 0.4748427, P(5) = 0.005063097.
2. The model estimates 3, P(3) = 0.573 to be the most likely chosen ordinal category when the stimulus gender is male (thus μ is -0.661 rather than 0). The likelihood estimates of the threshold coefficients are P(1) = 0.03026384; P(2) = 0.4065759; P(3) = 0.5733366; P(4) = 0.4792936 and P(5) 0.0006122306.
3. The middle – ordinal value 3 – is most often selected, as predicted by the model, but only for male and not for female stimulus gender. Nevertheless, it has usually a high probability. If the rationale is “that in Likert scale, the middle response (i.e., 3) should be removed and scales should be even because otherwise people will predominantly go for the middle, non-committal response, and the results will be useless”, the fact that model estimates 3 to be the most common value chosen for males, has usually a high likelihood of being selected, but is only the second most likely chosen value for females does only partially (if at all) support this claim. In addition, the model does only estimate the most likely selected values based on the fixed and random effect structure (approximating the data). This does not necessarily reflect the frequency (or percentage) of the values chosen in the data set; an inspection with the frequency table reveals that the mid value is not the most common chosen value in the total data set and in a spitted data set (according to male and female stimulus gender). But what does the one model fit for one experiment really mean on the grand scale when discussing the use of traditional Likert-scale (incl. midpoint) or forced choice scales?

Personally, I think the mid value probably will reflect, at least to some extent, a non-commitment bias. However, the true value of the mid value can also express a truly neutral opinion between disagreement on one side and agreement on the other for the participant.

On one hand, it is true that respondents might select a midpoint even if their true opinion is not neutral or might be used as a “dumping ground” in case of item unfamiliarity, items that are ambiguous (allow multiple interpretations or the participants rating is influenced by other confounds (e.g. social desirability; Stachowski, 2009; Nadler, Weston & Voyles, 2015). This is especially problematic from a data analysis perspective allowing valid interpretation of the data.

On the other hand, mainly from a test-theoretical perspective, including a midpoint gives respondents the opportunity to express a neutral opinion, especially on obscure topics (Johns, 2005) and respondents are not forced to agree or disagree, which may produce biased data. Eliminating the possibility to express a neutral opinion if the participant really has a true neutral opinion does compromise the validity of the interpretation in similar ways. A midpoint can also improve psychometric properties such as instrument reliability (Adelson & McCoach, 2010).

Thus, the more important question is not whether or not to include a midpoint, but rather when to omit or present a midpoint in a Likert-type scale (Chyung, Roberts, Swanson & Hankinson, 2017). The potential misuse of a midpoint as an ‘N/A proxy’ or ‘dumping ground’ can be improved by first improving clarity of survey items (Kulas & Stachowski, 2013) to avoid ambiguity and presenting other options off the scale such as Not applicable, I don’t know, or It depends (Kulas, et al., 2008; Raaijmakers, et al., 2000) to deal with situations where respondents are under strong social desirability pressures (John, 2005), or when it is expected that some respondents have little or no involvement in the survey topic (Weems & Onwuegbyzie, 2001).

Generally, I think it is ill-advised to a posteriori remove collected datapoints (mid-value), due to the following reasons: Removing the all data points containing mid values (1) will reduce the amount of observations in the data (N), (2) result in unnecessary and eventually costly collection of later undesired data and (3) might introduce systematic bias to the data.