CAMBRIDGE
UNIVERSITY PRESS

Research Article

The Galaxy Zoo catalogues for Galaxy And Mass Assembly (GAMA) survey

Benne Willem Holwerda[1] , Clayton Robertson[1], Kyle Cook[1], Kevin Pimbblet[2], Sarah Casura[3], Anne E. Sansom[4], Divya Patel[1], Trevor Alexander Butrum[1] , David Henry William Glass[4], Lee S. Kelvin[5,6,7] , Ivan K. Baldry[8], Roberto De Propris[9,10], Steven Bamford[11], Karen Masters[12], Maria Babakhanyan Stone[9] , Tim Hardin[1], Mike Walmsley[13], Jochen Liske[3] and S.M. Rafee Adnan[1]

[1]Department of Physics and Astronomy, University of Louisville, Louisville, KY, USA, [2]E.A. Milne Centre for Astrophysics, University of Hull, Kingston-upon-Hull, UK, [3]Hamburger Sternwarte, Universität Hamburg, Hamburg, Germany, [4]Jeremiah Horrocks Institute, University of Central Lancashire, Preston, UK, [5]Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA, [6]Department of Physics, University of California, Davis, CA, USA, [7]Astrophysics Research Institute, Liverpool John Moores University, IC2, LSP, Liverpool, UK, [8]Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool, UK, [9]FINCA, University of Turku, Turku, Finland, [10]Department of Physics and Astronomy, Botswana International University of Science and Technology, Palapye, Botswana, [11]School of Physics and Astronomy, University of Nottingham, Nottingham, UK, [12]Departments of Physics and Astronomy, Haverford College, Ardmore, PA, USA and [13]Dunlap Institute for Astronomy and Astrophysics, University of Toronto, Toronto, ON, Canada

Abstract

Galaxy Zoo is an online project to classify morphological features in extra-galactic imaging surveys with public voting. In this paper, we compare the classifications made for two different surveys, the Dark Energy Spectroscopic Instrument (DESI) imaging survey and a part of the Kilo-Degree Survey (KiDS), in the equatorial fields of the Galaxy And Mass Assembly (GAMA) survey. Our aim is to cross-validate and compare the classifications based on different imaging quality and depth. We find that generally the voting agrees globally but with substantial scatter, that is, substantial differences for individual galaxies. There is a notable higher voting fraction in favour of 'smooth' galaxies in the DESI+ zoobot classifications, most likely due to the difference between imaging depth. DESI imaging is shallower and slightly lower resolution than KiDS and the Galaxy Zoo images do not reveal details such as disc features and thus are missed in the zoobot training sample. We check against expert visual classifications and find good agreement with KiDS-based Galaxy Zoo voting. We reproduce the results from Porter-Temple+ (2022), on the dependence of stellar mass, star formation, and specific star formation on the number of spiral arms. This shows that once corrected for redshift, the DESI Galaxy Zoo and KiDS Galaxy Zoo classifications agree well on population properties. The zoobot cross-validation increases confidence in its ability to compliment Galaxy Zoo classifications and its ability for transfer learning across surveys.

Keywords: Galaxies: structure; galaxies: statistics; galaxies: spiral; galaxies: elliptical and lenticular; cD; galaxies: bulges

(Received 13 August 2024; revised 19 October 2024; accepted 26 October 2024)

1. Introduction

Identifying galaxy morphological features remains something where the human eye is unsurpassed. Despite great improvements in machine learning techniques, the subtleties of various galaxy morphological properties remain best identified by experts or groups of people. Arguably, this makes these classifications prone to human error and biases. To remedy biases and scale up to the kind of imaging surveys now practical, the Galaxy Zoo project was conceived (Lintott et al. 2008; Willett et al. 2013). By involving the wider public in classifying galaxies with a series of specific questions, one can both gain morphological classifications and estimate their biases and uncertainties.

Galaxy Zoo has had phenomenal success based on the work of thousands of volunteers working in tandem with professional astronomers. The results range from those that challenge our understanding of galaxy formation (e.g. Masters et al. 2011; Smethurst et al. 2015; Smethurst et al. 2022; Walmsley et al. 2019; Walmsley et al. 2021a; Walmsley et al. 2021b) to a wealth of rare objects (e.g. Keel et al. 2013; Keel et al. 2022; Keel et al. 2023).

Morphological classifications are useful on their own but increase dramatically in utility when combined with other information, specifically Spectral Energy Distribution (SED) fits to multi-wavelength data and spectroscopic redshifts. The Galaxy and Mass Assembly (GAMA, Driver et al. 2009; Driver 2021) survey is an excellent example of such a possible use. By combining MAGPHYS SED information from GAMA with specific Galaxy Zoo classifications for the survey area, subtle differences in specific star formation rates with a morphological feature can be identified (e.g. Porter-Temple et al. 2022; Porter et al. 2023; Smith et al. 2021). The improvements in accuracy, in both morphology from voting and inferred properties from the SED fit, for a statistical
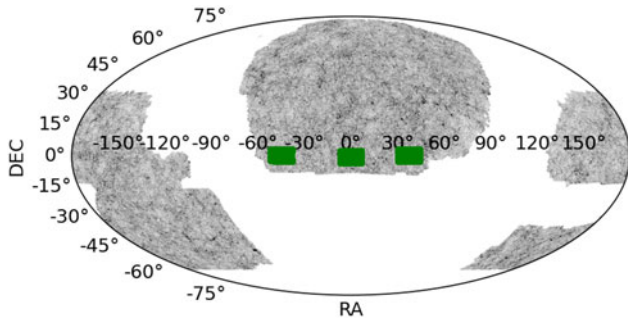
Figure 1. The relative coverage of the DESI-based Galaxy Zoo (black points Walmsley et al. 2023a) and the Galaxy And Mass Assembly (GAMA) equatorial fields (green) in a Molleweide projection. There is full coverage in both DESI and KiDS for the equatorial GAMA fields.

and complete samples allow for such more subtle relations to be revealed.

Our motivation to construct a GAMA Galaxy Zoo catalogue is therefore threefold: (a) to further explore morphology relations with other inferred properties for this highly complete sample, (b) to cross-examine morphology measures in two iterations of the Galaxy Zoo based on two different imaging surveys of the same galaxies (GAMA+KiDS and DESI Legacy Survey, Fig. 1) but completely different approaches (direct votes vs voting+ZOOBOT machine learning predictions), and c) provide a training catalogue for machine learning experiments for both researchers and students.

The GAMA-KiDS Galaxy Zoo (KiDS-GZ) and the DESI-LS Galaxy Zoo (DESI-GZ) classifications are fortuitously based on mostly the same classification flow diagram for Galaxy Zoo Dr4 (Fig. 2), with the number of options for three questions changing slightly. This means the same questions were answered by the volunteers but based on different images. An additional difference between the two Galaxy Zoo catalogues is that the DESI-LS classifications catalogue is that it is based on ZOOBOT machine learning classifications trained on the Galaxy Zoo classifications of all of DESI LS imaging. This study is meant to be a useful cross check between raw voting and a machine-learning extrapolated one.

This paper is organised as follows: in Section 2, we discuss the origin of the KiDS and DESI surveys Galaxy Zoo databases, as well as the parent GAMA survey catalogue. Section 3 briefly describes the differences in input imaging data quality from these surveys. In Section 4 we directly compare the reported voting fractions for each of the questions in Fig. 2, discussing patterns and their possible explanations. Section 6 uses the DESI Galaxy Zoo voting matched with the GAMA MAGPHYS catalogue to reproduce the results from Porter-Temple et al. (2022) as an test of how well DESI-GZ can be used in GAMA. In Section 7, we discuss the benefits and drawbacks of using either Galaxy Zoo catalogue for future morphological studies, and Section 8 lists our conclusions from this comparison.

2. Data

We use two iterations of the Galaxy Zoo catalogues: one based on KiDS images, made for the GAMA collaboration, and one made based on the DESI imaging survey as a larger Galaxy Zoo effort and then extrapolated using ZOOBOT (Walmsley et al. 2023a). The

relative coverage of these surveys, DESI and GAMA, are shown in Fig. 1. Ancillary data such as stellar mass estimates and redshifts are from the public GAMA-IV catalogues (DR4 Driver et al. 2022). The other catalogues (target and MAGPHYS) are part of GAMA DR2 release and available through the GAMA website.

2.1 GAMA-KiDS Galaxy Zoo

For GAMA, we use three sources: the GAMA survey itself, the KiDS survey for imaging, and finally the catalogue with voting from the Galaxy Zoo.

2.1.1 GAMA

GAMA is a combined spectroscopic and multi-wavelength imaging survey designed to study spatial structure in the nearby ($z < 0.25$) Universe on kpc to Mpc scales (see Driver et al. 2009, 2022, for an overview). The survey, after completion of phase 2 (Liske et al. 2015), consists of three equatorial regions each spanning 5 deg in Dec and 12 deg in RA, centred in RA at approximately 9h (G09), 12h (G12) and 14.5h (G15) and two Southern fields, at 02h (G02) and 23h (G23). The three equatorial regions, amounting to a total sky area of 180 deg^2, were selected for this study. For the purpose of visual classification, 49 851 galaxies were selected from the equatorial fields with redshifts $z < 0.15$ (see below). The GAMA survey is >95% redshift complete to $r < 19.8$ mag in all three equatorial regions (Driver et al. 2022). We use the MAG-PHYS SED fits data-products (Driver et al. 2018) from the third GAMA data-release (DR3, Baldry et al. 2018). The GAMA Galaxy Zoo voting catalogue is slated to be part of the rolling DR4 (Driver et al. 2022).

2.1.2 KiDS

The Kilo Degree Survey (KiDS, de Jong et al. 2013; de Jong et al. 2015; de Jong et al. 2017; Kuijken et al. 2019) is an optical wide-field imaging survey with the OmegaCAM camera at the VLT Survey Telescope. It has imaged 1 350 deg^2 in four filters (u g r i). The core science driver is mapping the large-scale matter distribution in the Universe, using weak lensing shear and photometric redshift measurements. Further science cases include galaxy evolution, Milky Way structure, detection of high-redshift clusters, and finding rare sources such as strong lenses and quasars. KiDS image quality is typically 0.″7 resolution (for sdss-r) and depths of 23.5, 25, 25.2, and 24.2 magnitude for i, r, g, and u, respectively. This imaging was the input for the Galaxy Zoo citizen science classifications (see also Kelvin et al. 2018).

2.1.3 KiDS Galaxy Zoo

Information on galaxy morphology is based on the GAMA-KiDS Galaxy Zoo classification (Lintott et al. 2008; Kelvin et al. 2018). GAMA-KiDS Galaxy Zoo data was initially provided by Lee Kelvin and Steven Bamford (private communication, 2019). Further details of this project may be found in Holwerda et al. (2019), Porter-Temple et al. (2022), and Porter et al. (2023). Briefly, Red-Green-Blue (RGB) cutouts were constructed from KiDS g-band and r-band imaging with the green channel as the mean of these. KiDS cutouts were introduced to the classification pool and mixed in with the ongoing classification efforts. For the Galaxy Zoo classification, 49,851 galaxies were selected from the equatorial fields with redshifts $z < 0.15$. Galaxy Zoo provided a monumental effort with almost 2 million classifications received from over 20 000 unique users over the course of 12 months. This classification has
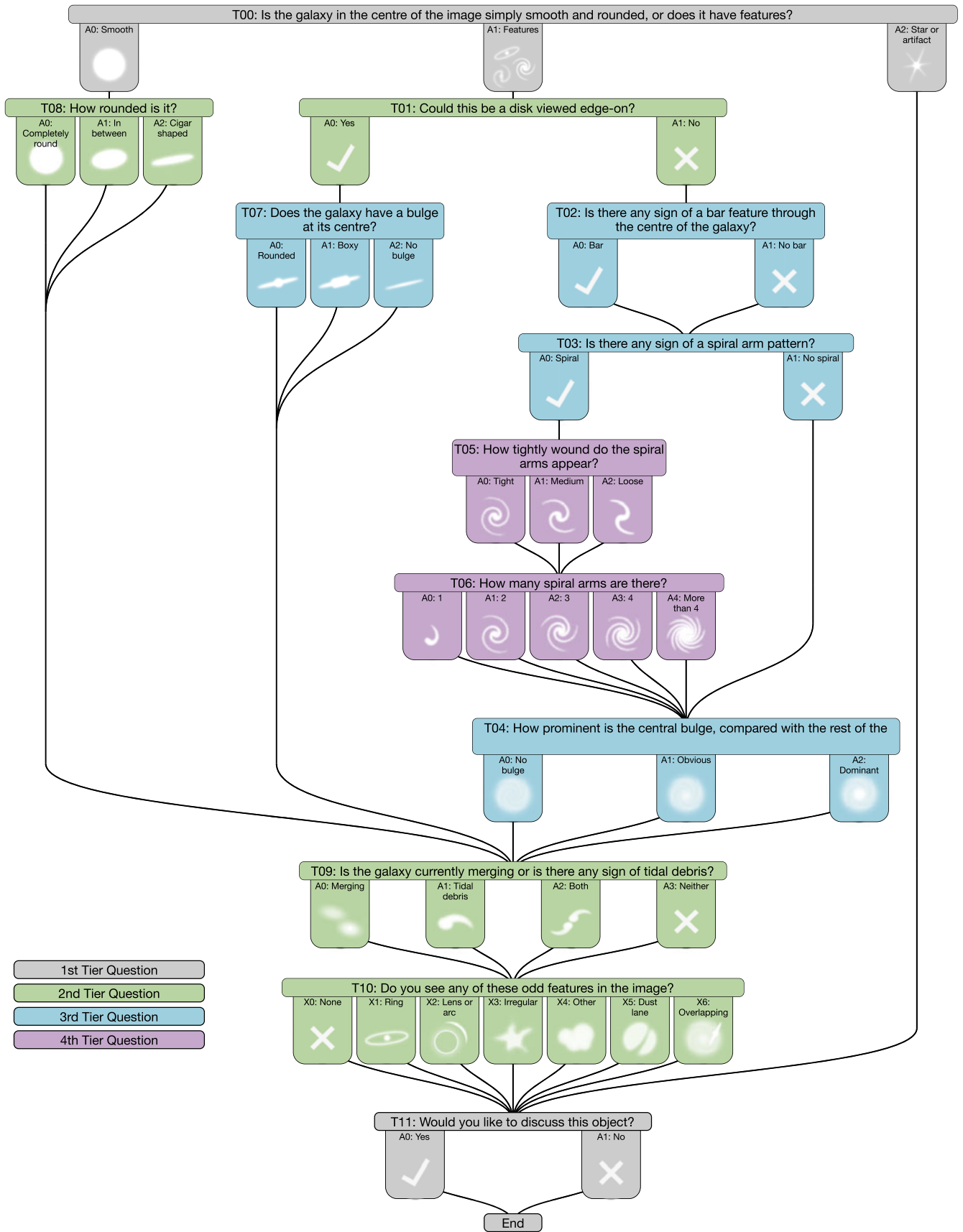
Figure 2. The flowchart of questions for volunteers for the KiDS-GZ database. Questions are listed in Table 2.

been used by the GAMA team to identify dust lanes in edge-on galaxies (Holwerda et al. 2019), searches for strong lensing galaxy pairs (Knabel et al. 2020), the morphology of green valley galaxies (Smith, Giroux, & Struck 2022a) and void galaxies (Porter et al. 2023) and to link star formation properties to the number of spiral arms (Porter-Temple et al. 2022). In this paper we use the visual classifications of disc galaxies from the Galaxy Zoo project; the full decision tree for the GAMA-KiDS Galaxy Zoo project is shown in Fig. 2. When talking about this catalogue, we use KiDS-GZ.

2.2 DESI Galaxy Zoo

Walmsley et al. (2023a) presents the galaxy zoo classifications based on three colour imaging of the Dark Energy Spectroscopic Instrument (DESI) survey (Dey et al. 2019). DESI requires images to target its spectroscopic fibres; these are primarily provided by the DESI Legacy Surveys (DESI-LS). These images are shallower than KiDS and often lower resolution (poorer seeing) which varied from field to field (Dey et al. 2019). The images from DESI-LS were converted into Red-Green-Blue (RGB) images for processing in the Galaxy Zoo classification pipeline.8

Walmsley et al. (2023a) provide user-friendly catalogues for use in future endeavours such as these. There are two catalogues: one with the Galaxy Zoo classifications over all of DESI-LS. These were used to train ZOOBOT (Walmsley et al. 2023b), a deep-learning tool, to predict voting fractions for all of DESI-LS sources. This is the catalogue shown in Fig. 1. Because DESI-LS covers a wide area, the number of galaxies with actual voting in the GAMA footprint is small. We use the ZOOBOT predictions catalogue for this reason to directly compare to the KiDS-GZ voting.

The ZOOBOT (Walmsley et al. 2023b) is a deep-learning tool that can be trained (and retrained) to predict Galaxy Zoo volunteer voting fractions based on RGB and greyscale images. The code is publicly available here: https://github.com/mwalmsley/zoobot. ZOOBOT is based on the Keras Python function *EfficientNetB0*, an image classifier using an optimised Convolutional Neural Network structure (Tan & Le 2019). The catalogues are publicly available here: https://zenodo.org/records/8360385. Because this is a prediction, not actual voting records for these galaxies, the voting fraction for example never reaches 100% in the DESI-GZ catalogue we are using (see section 4.1 on question T00).

2.3 Catalogue matching

We match the catalogues from KiDS-GZ and the DESI-GZ using their positions (the MATCH_COORDS_SKY algorithm in ASTROPY) within a match radius of two arcseconds (the width of the GAMA spectroscopic fibre). We opt to match only the GAMA equatorial fields (G09, G12, G15) and not G02 which also has Galaxy Zoo voting information because the equatorial fields have the most ancillary catalogues in the GAMA Data-releases (e.g. MAGPHYS SED fits). Previous work using the KiDS-GZ (Holwerda et al. 2019; Knabel et al. 2020; Knabel et al. 2023; Porter-Temple et al. 2022; Porter et al. 2023) all use only the equatorial fields.

Matching the GAMA targeting catalogue and DESI-GZ results in some 35k sources in common. The G02 field has 10k sources in common but these are not used in the rest of the paper. The DESI-GZ catalogue matches to GAMA are released as part of this paper with GAMA CATAID added for future GAMA archival use.

Table 1. Survey Depths of KiDS and the DESI Legacy Survey. The depth of DESI-LS depends strongly on which sub-survey is used but it is at least a magnitude shallower than KiDS.

	KiDS	DESI legacy survey
sdss-g	25.2	23.48–23.72
sdss-r	25	23.27–22.87
sdss-i	23.5	22.22–22.29
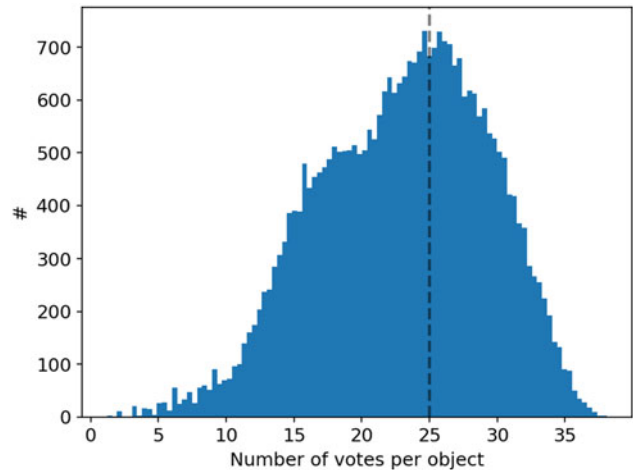
Figure 3. The number of votes in question T00 (see Fig. 2), which is asked for every galaxy. The mode for KiDS-GZ is 25 classifications per object (dashed line) and the mean 23.

3. Data quality

3.1 Image quality

The KiDS-GZ classifications are based on RGB images with the green filter constructed from red and blue, as the green filter (sdss-r) had not been observed for all targets yet at the beginning of the Galaxy Zoo iteration. The DESI-LS survey was obtained on a variety of telescopes but with the same filter set as SDSS and KiDS. Survey depths are summarised in Table 1. Spatial resolution varies similarly to depth (0.″6–1.″2) for DESI-LS while KiDS is more consistent with an average PSF of 0.″7. On the whole, the DESI-GZ classifications are based on more complete information, for example, including a green filter, and more area, but these are extrapolated using ZOOBOT. The GAMA-KiDS imaging is higher resolution and deeper in the two filters used.

The physical resolution in the respective surveys depends on the seeing and distance. In previous work (Smith et al. 2022b; Porter-Temple et al. 2022; Porter et al. 2023; Holwerda et al. 2022), samples were limited in redshift to keep sensitivity to smaller morphological features the same across the studied sample. In the case of the KiDS-GZ for example, a redshift limit of $z = 0.08$ corresponds to a physical resolution of ∼1 kpc, a key resolution for morphology (Lotz, Primack, & Madau 2004).

3.2 Number of classifiers

The histogram of the number of classifiers for each KiDS-GZ is shown in Fig. 3. The mean number is 36 classifiers for the

Table 2. The questions in the Galaxy Zoo 4th iteration (KiDS-GZ and DESI-GZ). The number of options are given. The number of options for T02, T04, and T06 were higher for DESI-GZ.

		# Options	
Number	Question	KiDS	DESI
T00:	Is the galaxy in the centre of the image simply smooth and rounded, or does it have features?	3	
T01:	Could this be a disc viewed edge-on?	2	
T02:	Is there any sign of a bar feature through the centre of the galaxy?	2	3
T03:	Is there any sign of a spiral arm pattern?	2	
T04:	How prominent is the central bulge, compared with the rest of the galaxy?	3	5
T05:	How tightly wound do the spiral arms appear?	3	
T06:	How many spiral arms are there?	5	6
T07:	Does the galaxy have a bulge at its centre?	3	
T08:	How rounded is it (the galaxy)?	3	
T09:	Is the galaxy currently merging or is there any sign of tidal debris?	4	
T10:	Do you see any of these odd features in the image?	7	

DESI-GZ training sample. The number of classifiers is lower for KiDS-GZ, leading to possibly larger scatter in the voting fractions. We note that the DESI-GZ catalogue, the KiDS-GZ is compared against is the extrapolated one by ZOOBOT.

The trade off between the two Galaxy Zoo iterations is that the KiDS-GZ imaging may be deeper and likely higher resolution, but the DESI-GZ has a higher number of votes going into individual sources of the large training sample which is then generalised to all the DESI-GZ catalogue (Walmsley et al. 2023a). If the *zoobot* classifier is trained well on the DESI-GZ, the accuracy should be equal to or surpassing the lower number of classifiers on the deeper KiDS data.

4. Voting comparison

Fig. 2 shows the flowchart of questions for the GAMA-KiDS classifications. We will compare the fraction of votes for each question.

Fig. 4 shows two KiDS images of ETGs examined by KiDS-GZ classifiers, and histograms of probabilities (multiples of vote fractions within relevant questions) for five endpoints within the flowchart in Fig. 2 (Glass 2024). The chosen endpoints, Smooth (T00:A0), Edge-On (T00:A1 × T01:A0), Spiral (T00:A1 × T01:A1 × T03:A0), No_Spiral (T00:A1 × T01:A1 × T03:A1) and Star/Artifact (T00:A2), form a complete set horizontally across Question T00 and probabilities therefore sum to 1. The threshold for morphology selection was the dominant probability above 0.4, to ensure a leading selection with only one other choice close behind, as is the case for GAMA64646.

T01-T03 are binary choices, allowing for easier comparisons using either fraction. In the case of multiple options, the voting fraction for each needs to be compared. We note that DESI-GZ

and KiDS-GZ had a different number of options for T02 and T04, where the DESI-GZ questionnaire had more options. In these cases, we compare the same worded answer, with much of the difference due to a difference in choice. We compare voting fractions since the number of votes is different for KiDS-GZ and DESI-GZ as the retirement criterion – the point in voting where the image was not shown to new classifiers – was set differently between these. Most usage focuses on voting fractions to identify features, not absolute or calibrated numbers of votes.

4.1 T00: Is the galaxy in the centre of the image simply smooth and rounded, or does it have features?

This is the first question encountered to separate those galaxies that are mostly featureless (elliptical/spheroidal) and artefacts from those with a lot of substructures (Fig. 2). This is a key question as the voter is not shown the remaining detailed questions if they do not mark the objects as 'having features'.

The voting in DESI-GZ does not reach 100% as does KiDS-GZ. This is an artefact of the ZOOBOT predictions. In the Galaxy Zoo voting in the DESI-LS, 100% fractions do happen.

Fig. 5 shows the comparison between the KiDS-GZ and the DESI-GZ ZOOBOT predicted voting fractions for the galaxies in common. This is one of the questions where we have answers for all the galaxies. T00 is the first question to be answered and thus always presented to all classifiers. A notable behaviour is that the fraction of positive votes for smooth in the DESI-GZ voting is typically higher than the KiDS-GZ one. This means Galaxy Zoo volunteers voted for galaxies to be smooth more in DESI-GZ than in GAMA-KiDS. This DEIS-GZ voting is then reflected in the ZOOBOT training sample. This is likely due to the deeper and/or higher resolution in the GAMA-KiDS images as volunteers are able to identify more galaxies with structure of some kind.

It also highlights a difference the ZOOBOT prediction makes on a voting fraction: there are no 100% voting fractions in the DESI-GZ. There are some in the KiDS-GZ but experience has shown these often to be the result of a single vote, especially for questions that depend on a previous choice (T01–T08). We leave the 100% fractions in our KiDS-GZ catalogue for now but often rejected these in analyses for this reason. An alternate approach is to renormalise these question with all the volunteers considering this object, not the number that answered the question.

4.2 T01: Could this be a disc viewed edge-on?

Rather than an axis ratio, this is a direct question if this is possibly an edge-on disc galaxy. As a binary question (only two answers possible; yes or no), we only need to examine one voting fraction because the other answer's listed fraction is the inverse of the first answer. This question immediately follows voting in favour of this galaxy having 'features'. The comparison sample is therefore smaller because a voting fraction is only recorded if someone in both groups of volunteers has voted on this galaxy to have features. There are two populations visible in Fig. 6, galaxies with a high fraction in both DESI-GZ ZOOBOT predictions and KiDS-GZ in favour of the edge-on perspective and a larger group with a small fraction ($f < 0.2$) in favour of the edge-on perspective.

The voter fractions in Fig. 6 agree quite well in case of the 'edge-on' question. This might make a sample of edge-on galaxies selected by this plot quite robust.
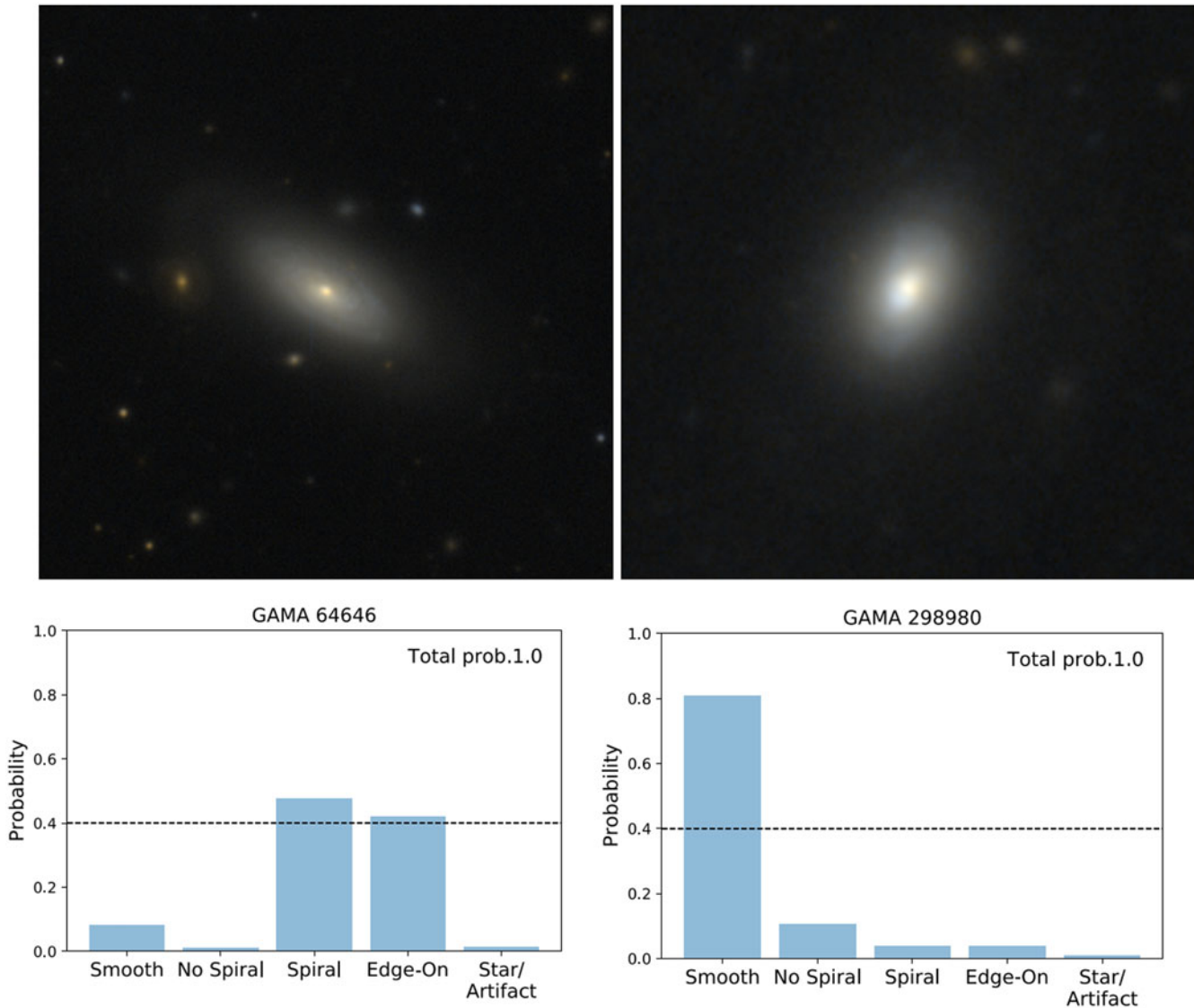
Figure 4. Two examples of KiDS galaxies with their respective voting fractions for questions T00, T01, and T03. From the thesis of David Henry William Glass (Glass 2024).

4.3 T02: Is there any sign of a bar feature through the centre of the galaxy?

This is a y/n question in the KiDS-GZ iteration, but one that was expanded to three options in DESI-GZ; none/weak/strong options for the bar. The fractions for 'no-bar' are compared in Fig. 7. This is the only option common to both surveys. There is reasonable agreement between the two GZ iterations but the change in options could mean that the fraction for no-bar in DESI is lower as the option for 'weak bar' is now available.

4.4 T03: Is there any sign of a spiral arm pattern?

The third and last binary (y/n) question. The fractions of voting are shown in Fig. 8. This is a question only answered after T00 and T01, and therefore the comparison sample is once again smaller. There is a good agreement for high fractions of voting in both DESI-GZ ZOOBOT predictions and KiDS-GZ.

Glass et al. (2024, in preparation. Glass 2024) showed that KiDS-GZ vote fractions can be used successfully to identify and remove weak spiral galaxies in samples of smooth early-type galaxies derived from GAMA classifications (Kelvin et al. 2014; Moffett et al. 2016). The GAMA classifications are based on SDSS imaging (catalogue VisualMorphologyv03), at lower resolution and depth than KiDS. For disc-like early-type galaxies (ellipticals) in their sample, 19% were found to have weak spiral features using KiDS-GZ.

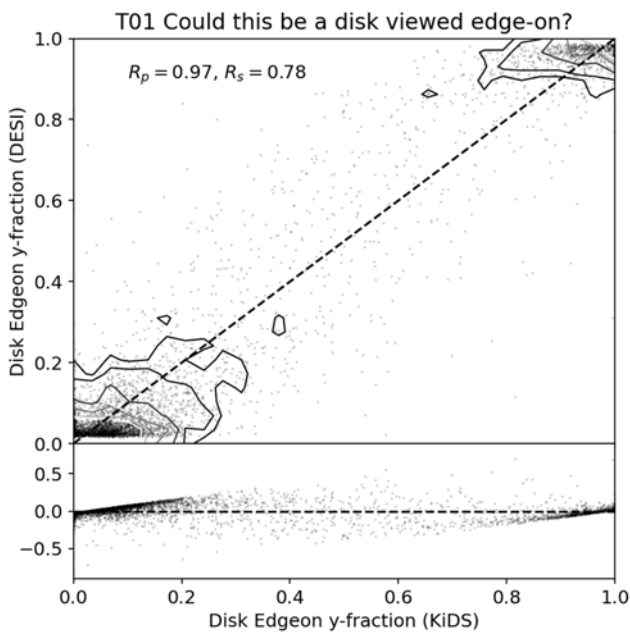### 4.5 T04: How prominent is the central bulge, compared with the rest of the galaxy?

This is a question with three options in the KiDS-GZ but five options in the DESI-GZ and the subsequent ZOOBOT predictions. In principle, it could be reduced to a binary one (evidence of a bulge y/n?) similar to T07.
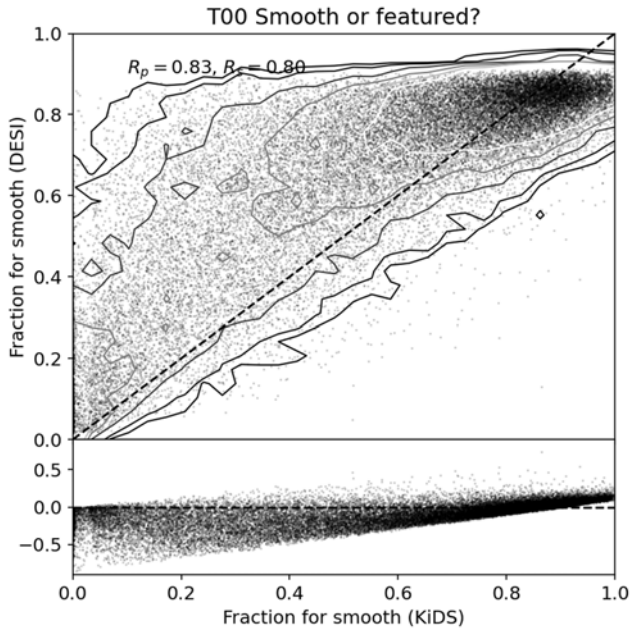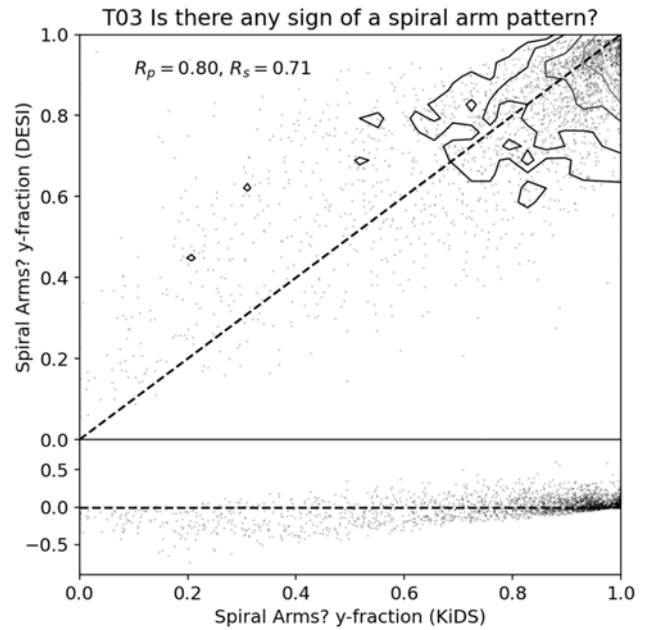
Figure 5. The fractions of votes in question T00 (Table 2) in favour of these galaxies to be 'smooth'. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ voting fraction predictions by the ZOOBOT (top panel) or the difference between the two (bottom panel). Each point is a galaxy in common between the DESI-GZ and the KiDS-GZ. T00 is one of three questions that is asked for each object. Contours are drawn at 5,10,25,50,75, and 100 densities. R_P and R_S are the Pearson and Spearman ranking in the relation, respectively.

Figure 7. The fractions of votes in question T02 (Table 2) in favour of these galaxies to *not* have a bar. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel). This is one of the questions that changed between KiDS-GZ and DESI-GZ with the addition of 'strong' and 'weak' in the latter.

Figure 8. The fractions of votes in question T03 (Table 2) in favour of these galaxies to have an identifiable spiral pattern. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ voting fraction (top panel) or the difference between the two (bottom panel).
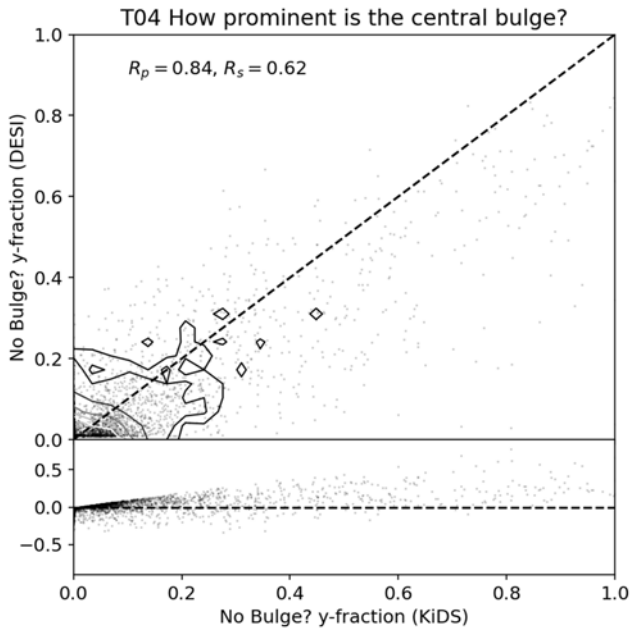

Figure 6. The fractions of votes in question T01 (Table 2) in favour of these galaxies to be 'edge-on'. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel). Values of exact 0 and exactly 1 are ignored as these likely indicate single (erroneous) vote counts.

Fig. 9 shows the 'no bulge' voting fractions for KiDS-GZ and DESI-GZ. This is the answer both Galaxy Zoo questions have in common. The agreement with other answers for this question is quite poor. The most likely explanation is that the DESI-GZ offered more options to classify. Many of the prominent bulges

Figure 9. The fractions of votes in question T04 (Table 2) in favour of these galaxies to have a 'prominent' central bulge. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel).

Figure 10. The fractions of votes in question T04 (Table 2) in favour of these galaxies to have a 'tightly wound' spiral arms. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel).

would have gotten votes for 'large' bulge instead. Other possible explanations include the difference in depth between DESI-GZ and KiDS-GZ and that the RGB images were constructed differently in each survey. The fact that 'no bulge' is consistent is encouraging.

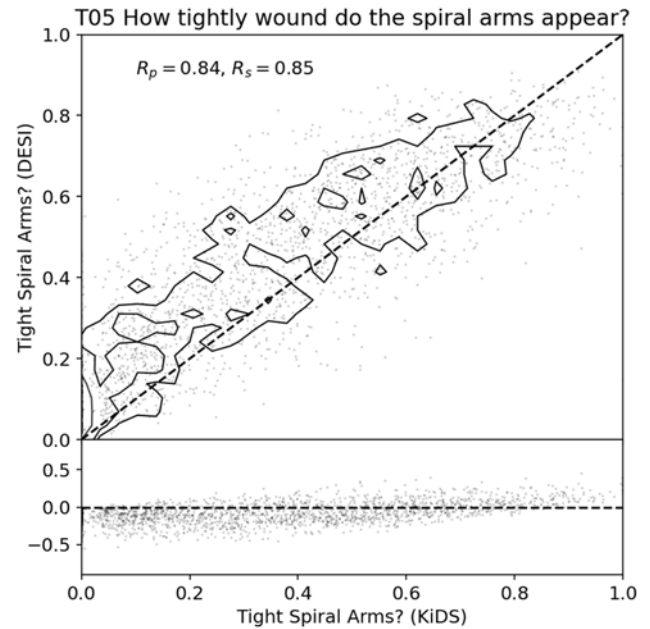### 4.6 T05: How tightly wound do the spiral arms appear?

This question is only presented if the Galaxy Zoo volunteer answers affirmative to T03. The three answers are not easily reduced to a binary question. Fig. 10 shows the voting fractions for the 'tightly wound' answer compared between KiDS-GZ and DESI-GZ ZOOBOT predictions.

The reasonably good agreement between the answer fractions for KiDS-GZ and DESI-GZ ZOOBOT predictions is reassuring that once spiral arm features are identified in either survey, there is good agreement on their appearance.

4.7 T06: How many spiral arms are there?

This is the question on which Porter-Temple et al. (2022) focused for their study of star formation and stellar mass properties. The option 'more than 4' for the number of spiral arms is functionally a vote for a flocculent spiral. In principle, this question could be reduced to a binary one of 'grand design spiral' or 'flocculent' but this has not been used. Porter-Temple et al. (2022) found the vast majority of objects in their sample ($\log(M_*/M_\odot) > 9$, $z < 0.08$) for completeness and resolution reasons) to mostly consist of 2-armed spirals with much smaller but statistically significant numbers of the other categories.

Fig. 11 shows the voting fractions for KiDS-GZ and DESI-GZ ZOOBOT predictions for all four distinct answers in question T06. The best agreement is for 2-armed spirals (the most common kind). One-armed spirals are agreed upon with a low fraction,

and a similar but more diffuse version of that pattern repeats for 3-armed spirals. There is a noticeable trend with both 3- and 4-armed spirals where there is a higher voting fraction in favour of these in the KiDS-GZ compared to the DESI-GZ ZOOBOT predictions. There is an additional answer possible in DESI-GZ ('cannot tell') which was not an option in KiDS-GZ. This all suggests the threshold for identifying a larger number of arms may have to be set to a lower fraction in the DESI-GZ ZOOBOT predictions. There are different ways to ensure a galaxy has a certain number of spiral arms. One could require a simple majority (>20% for a given option within the five options in T06) or an overwhelming consensus (>50%), depending on how certain one wants to be of the selected sample. The former can still be a close call (all options have almost 1/5th of the vote), while the latter is unambiguous but lower statistics. This question is relevant for any comparison to Porter-Temple et al. (2022) or Hart et al. (2017): 'how well do volunteers agree on the number of spiral arms in a galaxy' (see 6).

4.8 T07: Does the galaxy have a bulge at its centre?

This question is only answered if T01 is positive (view is edge-on). Three answers are possible (Fig. 2). In principle, this question can be reduced to a binary one (is there a bulge y/n?) by combining the voting of the first two options ('boxy' and 'round'). Fig. 12 shows the fraction of votes in favour of 'no bulge'. Generally speaking, the two Galaxy Zoo iterations, KiDS-GZ and DESI-GZ ZOOBOT predictions are in broad agreement but with large scatter.

4.9 T08: How rounded is it?

This is the only dedicated question if the volunteer answers 'Smooth' in T00. This is exclusively for elliptical/spheroidal galaxies. Fig. 13 shows the voting fractions for all three options (round, in-between, and cigar) compared between the KiDS-GZ and
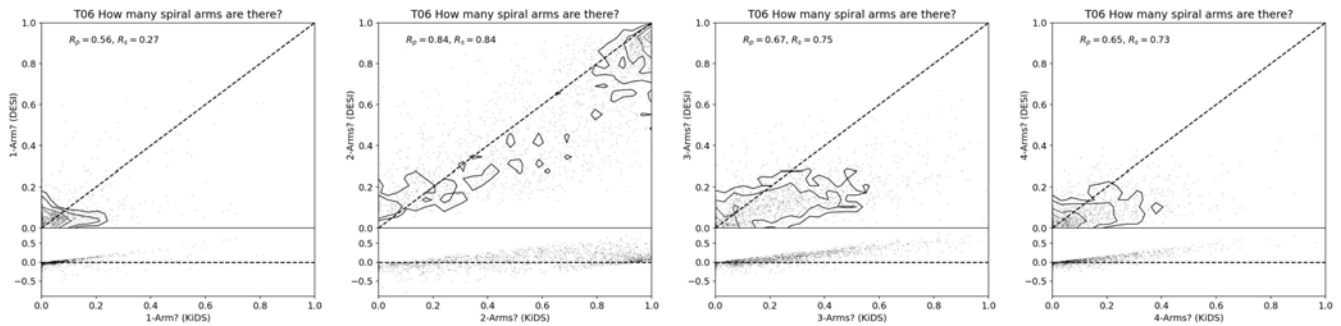
Figure 11. The fractions of votes in question T06 (Table 2) on the number of spiral arms, one to four from left to right. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel).
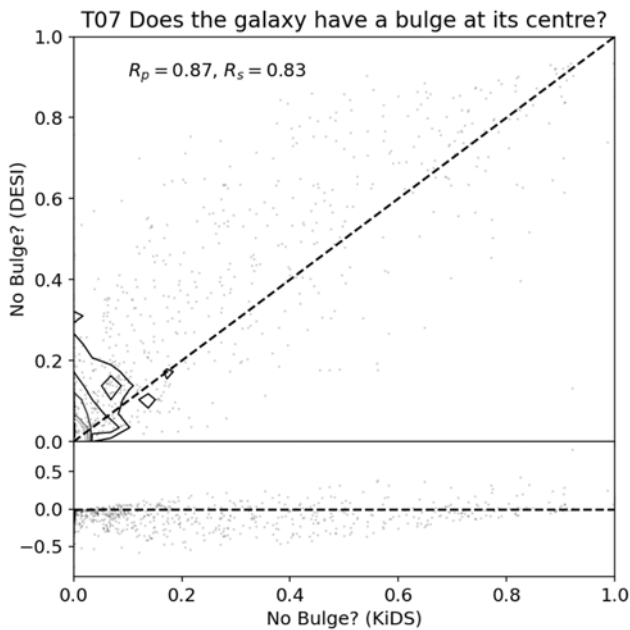
Figure 12. The fractions of votes in question T04 (Table 2) in favour of these edge-on galaxies to have no bulge. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel).

DESI-GZ ZOOBOT predictions. There is general agreement but with substantial scatter, to be expected for a slightly relative or subjective question.

4.10 T09: Is the galaxy currently merging or is there any sign of tidal debris?

This is a question with four possible answers but could be reduced to a single 'signs of interaction y/n?' by combining the first three answers to compare against the 'none' voting fraction. This was the usage in Porter et al. (2023) for void galaxies. The middle options between 'merging' and 'none' in this question are the only change between KiDS-GZ and DESI-GZ ZOOBOT predictions. The middle options are 'tidal debris' and 'both' (meaning the galaxies show both signs of merging and tidal debris) for the KiDS-GZ and the middle options in DESI-GZ ZOOBOT predictions are 'minor' and

'major' indicating the relative ratio of the galaxies in minor/major interaction.

Fig. 14 shows the comparison between the two answers the GZ iterations have in completely common (i.e. the answer is phrased the same). This is a question that is asked for all galaxies, regardless of T01 so the comparison has more statistics. Generally, there is a reasonable agreement (high fraction of no-merger in both iterations) but especially at lower fractions (more ambivalence), the scatter is higher. There is a better agreement on no-merger than on merging since the other options could draw votes away depending on the image depth (i.e. a tidal feature is visible in KiDS but not in DESI, either in the volunteer voting or the ZOOBOT predictions).

4.11 T10: Do you see any of these odd features in the image?

The final question is unique in that the answers are not mutually exclusive and one could vote for more than one of these features. For example, one could see an overlapping pair of galaxies and a prominent dust lane visible. Whether or not it is clear to each volunteer that multiple answers are allowed is not clear.

This question was not included in the data-release by Walmsley et al. (2023a), and we do not include the comparison here. The question was undoubtedly asked but it would be difficult to inter-compare with likely low statistics as these are relatively infrequently occurring phenomena. ZOOBOT predictions for these are difficult for the same reasons.

4.12 Correlation metrics

Figs. 5 through 14 include the Pearson (R_p) and Spearman (R_s) rankings. These are summarised in Table 3 including the p-values returned with each test. The Pearson ranking is an indication of how linear the relation between the two voting fractions is. The Spearman one is a ranking for a monotonous, but importantly not necessarily linear, relation between the two voting fractions.

Most of the voting between KiDS-GZ and DESI-GZ ZOOBOT predictions is highly correlated with rankings well above 0.8. The highest agreement is on whether this disc can be viewed edge-on (T01). This is reflected in Fig. 6 with clusters at 0 and 1.

The lower correlations are often for questions where either there were more options in one of the Galaxy Zoo iterations (e.g. T04) or a suspected dependence on surface brightness (e.g. T03), or both.

Of the number of spiral arms, a subtle difference in surface brightness may make a difference. The agreement is the strongest
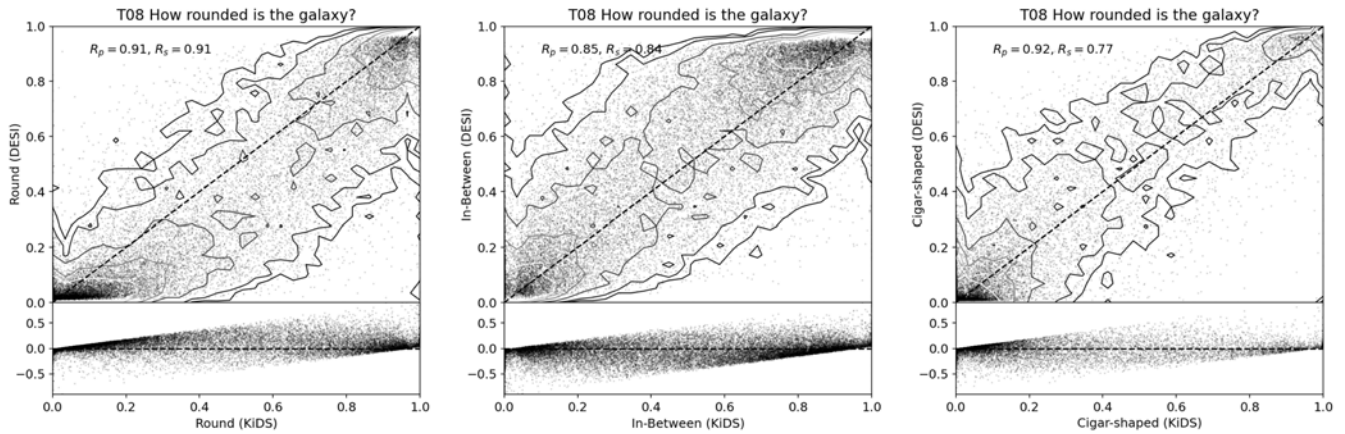
Figure 13. The fractions of votes in question T08 (Table 2) in favour of smooth galaxies that appear round (left panel), in-between (middle panel), or cigar-shaped (right panel). On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel).
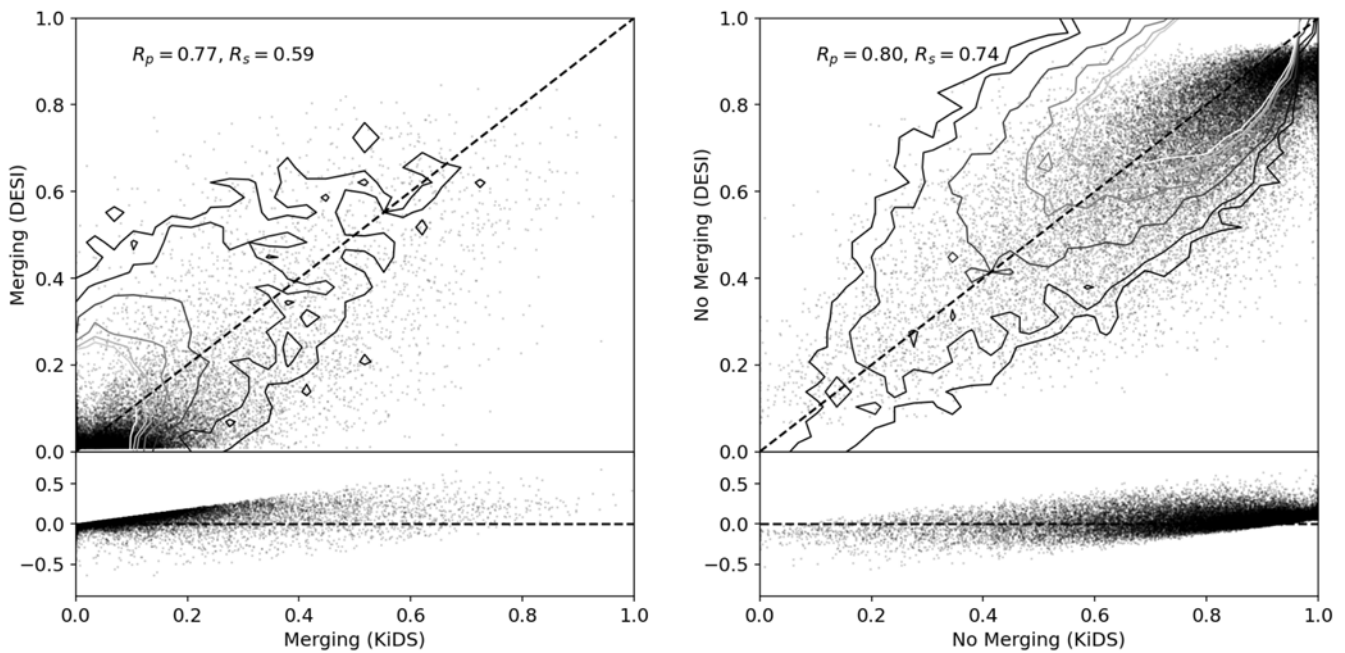
Figure 14. The fractions of votes in question T09 (Table 2) in favour of an interaction or not. On the *x*-axis is the voting fraction for the KiDS-GZ and on the *y*-axis the DESI-GZ ZOOBOT predicted voting fraction (top panel) or the difference between the two (bottom panel).

for two arms, where the statistics are the highest. The agreement on mergers (T09) are surprisingly good because these are dominated by the answer that there is no evidence for an ongoing or past merger (Fig. 14).

5. Other GAMA visual classifications

Previous visual classifications of the GAMA galaxies include those by the GAMA team (Driver et al. 2022) and low-redshift quasar hosts (Stone et al. 2023). These are visual classifications of the

galaxy as a whole (Table 4). We compare the KiDS-GZ voting against these expert visual classifications.

Fig. 15 shows the overlap (58 galaxies out of 205) with the DESI-GZ sample and the one from Stone et al. (2023) for quasar host galaxies and question T00, the most numerous and relevant one for the morphology of the galaxy as a whole. Early type galaxies (E or S0) show a low fraction of 'features' votes, later types, that is, disc-dominated classes have higher fractions of votes in favour of 'features'. This makes the Galaxy Zoo classifications consistent with the expert visual assessment in Stone et al. (2023) of quasar host galaxies.

Table 3. The questions in the Galaxy Zoo 4th iteration (GAMA-KiDS and DESI-LS ZOOBOT predictions). The number of options are given.

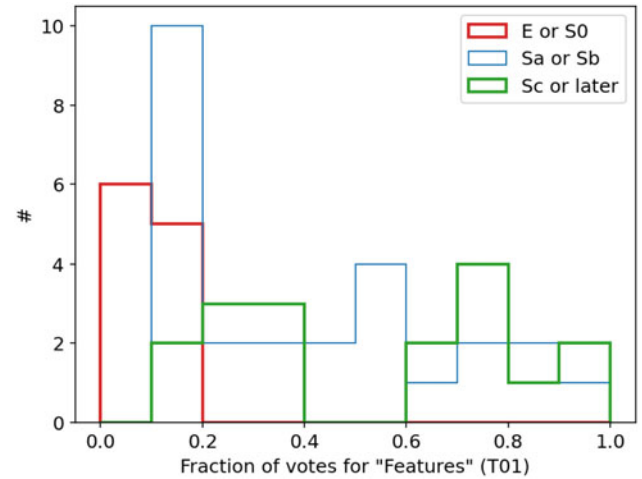| Question	Pearson R (p-value)	Spearman R (p-value)	Figure
T00	0.83 (0.00)	0.80 (0.00)	Fig. 5
T01	0.97 (0.00)	0.78 (0.00)	Fig. 6
T02	0.79 (0.00)	0.78 (0.00)	Fig. 7
T03	0.80 (0.00)	0.71 (0.00)	Fig. 8
T04	0.84 (0.00)	0.62 (0.00)	Fig. 9
T05	0.84 (0.00)	0.85 (0.00)	Fig. 10
T06-1	0.56 (1.48E-169)	0.27 (2.18E-35)	Fig. 11
T06-2	0.84 (0.00)	0.84 (0.00)	Fig. 11
T06-3	0.67 (1.28E-275)	0.75 (0.00)	Fig. 11
T06-4	0.65 (3.28E-248)	0.73 (0.00)	Fig. 11
T07	0.87 (0.00)	0.83 (1.58E-272)	Fig. 12
T08-1	0.91 (0.00)	0.91 (0.00)	Fig. 13
T08-2	0.85 (0.00)	0.91 (0.00)	Fig. 13
T08-3	0.92 (0.00)	0.91 (0.00)	Fig. 13
T09-1	0.77 (0.00)	0.59 (0.00)	Fig. 14
T09-2	0.80 (0.00)	0.74 (0.00)	Fig. 14

Table 4. Visual classification schemes used in Driver et al. (2022) and Stone et al. (2023) for GAMA galaxies.

NumClass	Class	Description
		Driver et al. (2022) classifications
0	E	Elliptical system with a single visual component
1	cBD	Two-component system with a compact high-surface brightness bulge
2	dBD	Two-component system with a diffuse or extended bulge (or bulge complex)
3	D	Disc system with a single visual component
		Stone et al. (2023) classifications
0	–	Early-type (E or S0)
1	–	Middle-type (Sa or Sb)
2	–	Late-type (Sc or later)
3	–	Unknown types (possible merger)

Figure 15. The histogram of classified galaxies as a function of the voting fraction of T00 question 'smooth or featured?' in the DESI-GZ. The Stone et al. (2023) classifications show a low fraction for ellipticals (E-S0) and higher voting fraction for disc galaxies (either 'Sa or Sb' and 'Sc or later'). This makes the Galaxy Zoo classifications consistent with the expert visual assessment from Stone et al. (2023).

Figure 16. The fraction of voting for T00 question 'smooth or featured' for the visual classifications presented in Driver et al. (2022): Elliptical, compact-bulge (cBD), diffuse disc (dBD), and disc-dominated (D) galaxies. Ellipticals have the lowest voting fractions, followed by disc-dominated and both cBD and dBD have a high voting fraction for 'features'.
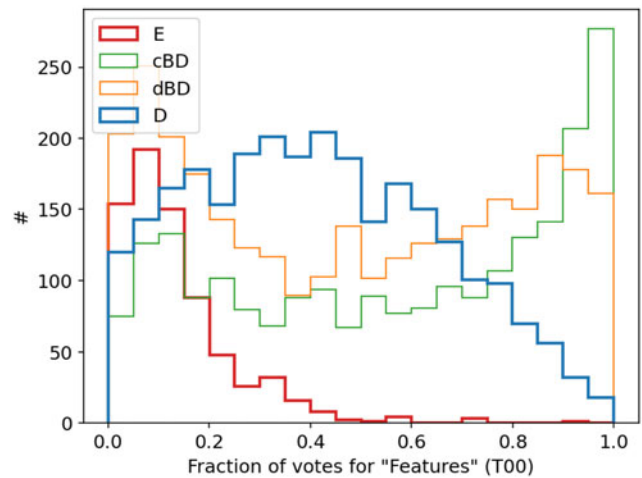
There is a larger sample of overlap between the visual classifications by Driver et al. (2022) and the KiDS-GZ catalogue. The classifications by Driver et al. (2022) focus on the prominence of the bulge with respect to the galaxy as a whole. Both T00 and T04 are therefore good comparison questions (Table 2, Fig. 2).

Fig. 16 shows the distribution of voting for the first question 'smooth or featured?'. Ellipticals have the lowest voting fractions, followed by disc-dominated classes (cBD, dBD and D). Both cBD and dBD have a high voting fraction for 'features', higher than disc alone (D).

The voting for T04 'bulge prominence?' in Fig. 17 for the answer 'no bulge'. The highest fraction is the 'pure disc' (D), which is consistent with the KiDS-GZ vote. The diffuse bulge (dBD) is next, followed by the concentrated bulge (cBD). Ellipticals are a

small fraction of the galaxies in this question as most have been filtered out by T00.

Overall the expert classifications and the KiDS-GZ classifications agree well. Figs. 16 and 17 can serve as a possible translation between Galaxy Zoo voting and expert classes (e.g. $0.2 < f_{T00} < 0.7$ and $f_{T04} > 0.5$ would select a fairly clean disc-only sample from the Galaxy Zoo voting.

6. Comparison to Porter-Temple+ (2022)

Using the KiDS-GZ, Porter-Temple et al. (2022) examined the dependence of stellar mass, star formation rate, and specific star formation rate with the number of spiral arms. They adopted
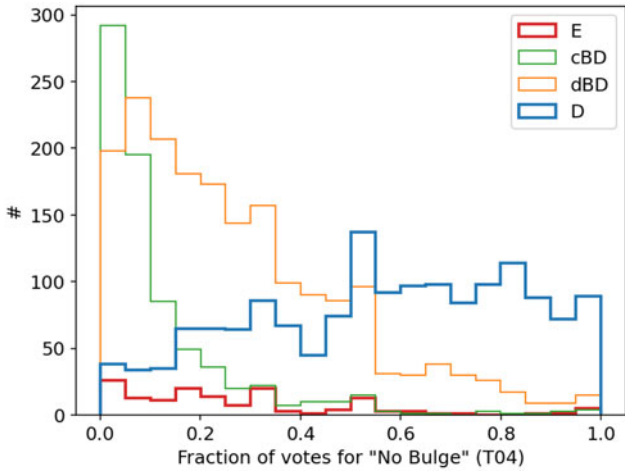
Figure 17. The 'no bulge' fraction of voting for T04 question 'How prominent is the bulge?' for the visual classifications presented in Driver et al. (2022): Elliptical, compact-bulge (cBD), diffuse disc (dBD) and disc-dominated (D) galaxies. This is the other question that can be directly compared to the Driver et al. (2022) classifications as these focus on the prominence of the bulge. Highest fraction is the 'pure disc' (D), followed by the diffuse and concentrated bulge classes. Ellipticals are rarely in this question.
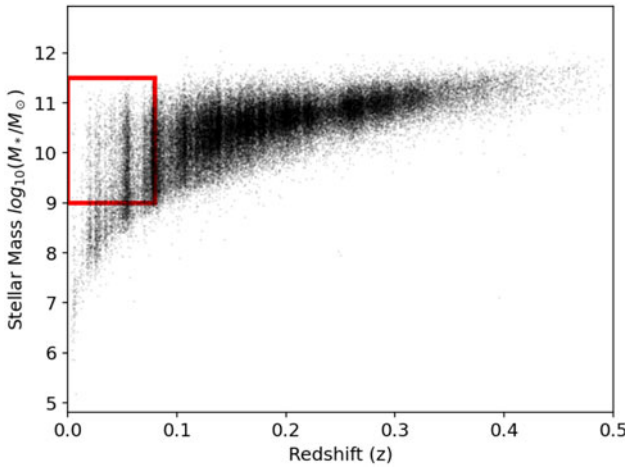

Figure 18. The redshift vs stellar mass as measured by MAGPHYS for the GAMA galaxies with DESI-GZ ZOOBOT predictions classifications. The red delineated area is the selection used in Porter-Temple et al. (2022). We select the same redshift range and stellar mass range for comparison but use the DESI-GZ ZOOBOT predictions voting for the classification on number of spiral arms.

a conservative approach in the identification of spiral arms by limiting the redshift to $z < 0.08$, adopting a lower limit of $log(M_*/M_\odot) = 9$, and setting a relatively high threshold for a galaxy to be classified with one, two, three, four, or five and more spiral arms ($f > 0.5$).

Their selection criterion is shown as the red line box in Fig. 18. The KiDS-GZ classifications were limited intentionally to $z = 0.15$ but we can see from the GAMA galaxies with DESI-GZ ZOOBOT predictions classifications in Fig. 18 that this limit is not enforced for DESI-GZ ZOOBOT predictions. This is not the reason there is a much higher fraction in voting for smooth galaxies in DESI-GZ ZOOBOT predictions (Fig. 5) because that sample is limited to $z =$
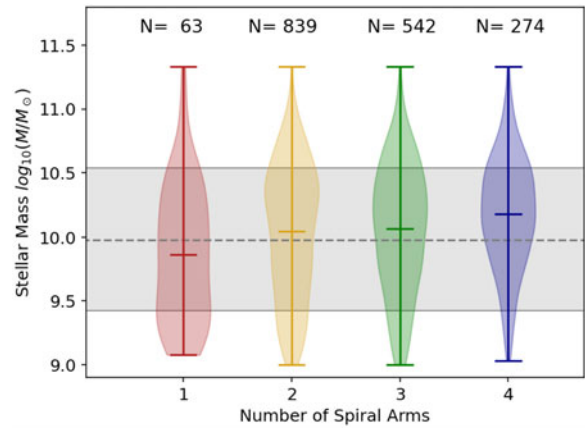
Figure 19. The distribution of stellar mass as measured by MAGPHYS vs the number of spiral arms for the GAMA galaxies with DESI-GZ ZOOBOT predictions classifications.
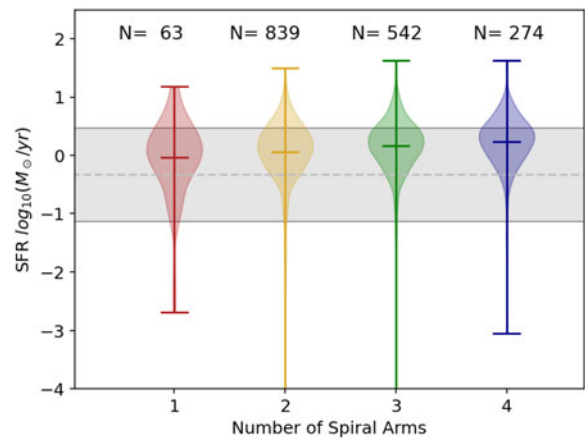
Figure 20. The distribution of star formation rate as measured by MAGPHYS vs the number of spiral arms for the GAMA galaxies with DESI-GZ ZOOBOT predictions classifications.
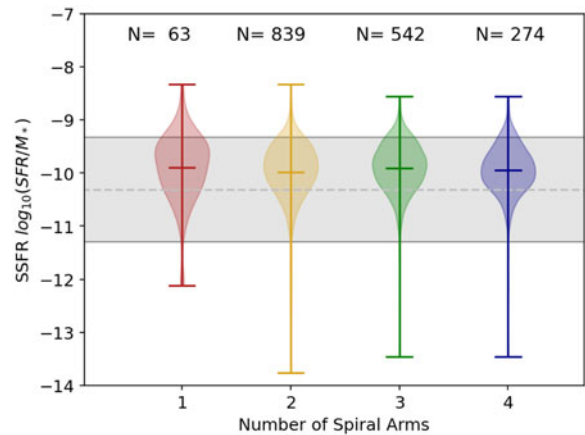
Figure 21. The distribution of the specific star formation rate as measured by MAGPHYS vs the number of spiral arms for the GAMA galaxies with DESI-GZ ZOOBOT predictions classifications.

Table 5. The Kolmogorov-Smirnov test and associated p-value for the number of arms (m). The comparison is between the population with m spiral arms and the full population. Stellar mass, specific star formation, and star formation between each spiral arm number and the total population in Fig. 18. Differences between the populations are similar in size and significance than they were in Porter-Temple et al. (2022) using the KiDS-GZ classifications. Their K-S and p-values (their Table 1) are reproduced next to each column.

m	Stellar Mass	P-T	SFR	P-T	sSFR	P-T
1	0.15 (0.1203)	0.138 (0.038)	0.21 (0.0063)	0.099 (0.256)	0.23 (0.0023)	0.234 (0.000)
2	0.08 (0.0011)	0.034 (0.203)	0.25 (0.0000)	0.110 (0.000)	0.17 (0.0000)	0.072 (0.000)
3	0.11 (0.0000)	0.152 (0.001)	0.34 (0.0000)	0.281 (0.000)	0.21 (0.0000)	0.187 (0.000)
4	0.22 (0.0000)	0.252 (0.081)	0.40 (0.0000)	0.291 (0.028)	0.22 (0.0000)	0.261 (0.064)

0.15 by the crossmatch with KiDS-GZ. The images on which DESI-GZ ZOOBOT predictions voting are based are shallower (Table 1) and thus more prone to miss lower surface brightness features (spiral arms, tidal arms etc).

Voting in favour of T06 options other than 2-arms show slightly lower fractions for the same galaxies compared to the KiDS-GZ (Fig. 11, reflected in lower rankings as well). Therefore, we adopt slightly less stringent criteria to classify a galaxy with a certain number of spiral arms: we require that the $f_{disc}(T00) > 0.3$ and $f_{n-arm}(T06) > 0.2$ with n the number of spiral arms. We also require the redshift to be $z < 0.08$ as the DESI imaging is not higher resolution than KiDS and a minimum stellar mass of $log_{10}(M_*/M_\odot) > 9$. The lower voting fraction than Porter-Temple et al. (2022) for a choice of n arms is needed because otherwise the statistics for any number other than $n = 2$ would be too low for a comparison. We note that 5+ still suffered from too low numbers to be included in the plot.

Fig. 19 shows the distribution of stellar masses for $n = 1$, 2, 3, or 4 spiral arms. The $n = 5+$ category did not get enough votes for a statistically significant result. We see a similar rise in stellar mass with the number of spiral arms as Porter-Temple et al. (2022), compare to their Fig. 4.

Fig. 20 shows the star formation rate of galaxies with $n = 1$, 2, 3, or 4 arms. Similar to Porter-Temple et al. (2022), their Figure 6, we see a rise in the star formation with number of spiral arms, similar to the increase with mass.

Fig. 21 shows the specific star formation rate (SFR/M_*) of galaxies with $n = 1$, 2, 3, or 4 arms. Similar to Porter-Temple et al. (2022), we see a flat or slight decline in the specific star formation with number of spiral arms. A very similar, subtle decline in sSFR with the number of arms was observed by Porter-Temple et al. (2022) in their Figure 8.

The comparison in stellar mass, star formation, and specific star formation can be done by comparing the distribution of values of galaxies with a certain number of arms (m) to the population as a whole. The similarity can be tested with the Kolmogorov-Smirnov test (K-S), which measures the greatest fractional difference in the cumulative distribution; 0 means no difference, and 1 means completely different distributions. The K-S test values are listed in Table 5 with the p-value in brackets. The differences in distributions are not very large, similar to what Porter-Temple et al. (2022) found and show the same trends. The two-armed spiral, being the most numerous, will resemble the population at large the most, with the lowest K-S value. The trend is higher K-S values away from 2-arms. These are all the same trends observed by Porter-Temple et al. (2022). We conclude that with accurate inferred parameters – stellar mass, star formation, and specific
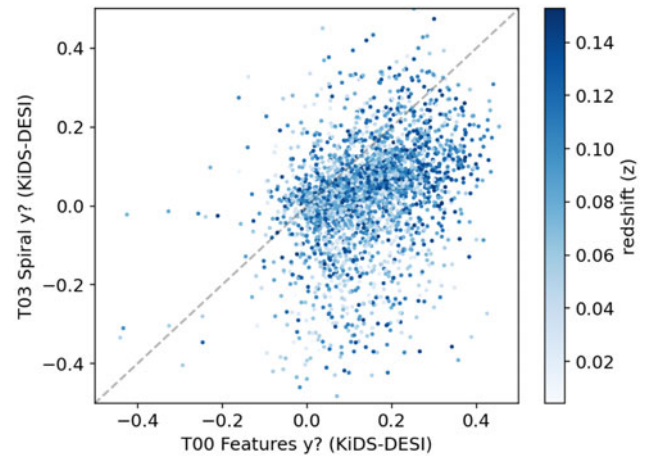
Figure 22. The difference in voting fraction between KiDS and DESI-based Galaxy Zoo on question T00 'smooth or features?' on the x-axis and T03 'Spiral Structure?' on the y-axis. The KiDS voting favours features over the DESI ZOOBOT predictions but spiral structure is identified with similar frequency. T00 difference is not distance dependent and must be an inherent difference between DESI ZOOBOT predictions and KiDS voting.

star formation rates – one can reproduce the experiment from Porter-Temple et al. (2022) accurately.

7. Discussion

The two Galaxy Zoo iterations agree reasonably well with each other despite different approaches to the imaging data that went into them and only slight differences in the classification questions. These were the same galaxies and observed in the same filters (*gri*), but on different telescopes, under different seeing conditions, to different depths, with a different approach to the generation of an RGB colour image, based on different numbers of classifiers, and in the case of DESI, extrapolated by ZOOBOT.

The correlation between answers (Table 3) show very good (linear) agreement between the voting fractions for most of the features. This adds to the confidence that these features are present in these galaxies, separately from the origin of the voting. There is excellent agreement between KiDS-GZ and DES-GZ ZOOBOT predictions.

The question whether a volunteer sees features (T00) and whether they see spiral structure (T03) are somewhat overlapping; one would need to see features or disc structure to even see spiral arms. It is therefore perhaps illustrative to compare the difference in voting fractions in DESI/ZOOBOT and KiDS-based Galaxy

Zoo. We do so in Fig. 22. There is a clear shift of voting in the KiDS-based Galaxy Zoo towards galaxies with features in T00 but then, once features have been found, the voting is mostly balanced around a difference of 0 for the T03. In the DESI ZOOBOT predictions, more galaxies are identified as 'smooth' but once features are identified, the result is similar to the KiDS one. This is not distance dependent and it is most likely the result of the depth of DESI compared to KiDS, influencing the final ZOOBOT predictions.

There is a lot of scatter in the fractions of votes. For individual galaxies, there may be room for interpretation, a well-known effect even among expert classifiers (cf discussion in Nair & Abraham 2010a; Nair & Abraham 2010b). But for statistical uses, either catalogue looks to agree well with one another. Apart from perhaps removing unity values in the voting fraction, not much more correction is needed in KiDS-GZ.

We check the KiDS-GZ voting with expert visual classifications by Stone et al. (2023) and Driver et al. (2022). Both agree in broad terms with the voting in the Galaxy Zoo catalogue. The distributions of voting fractions for T00 and T04 agree with the categories assigned by experts, for example, pure disc galaxies have a high voting fraction for no bulge and early types have a low voting fraction for no features. The broad agreement is another validation of the utility of this Galaxy Zoo catalogue for future uses.

It is heartening to see that previous results by Porter-Temple et al. (2022) are recovered here. The DESI-GZ ZOOBOT predictions has voting for higher redshift galaxies, and it is based on shallower imaging data, thus consistency with previous results strengthens its use case. Similarly, the KiDS-GZ voting fractions were not calibrated or de-biased but the higher thresholds compensate for that. Results like those in Figs. 19–21 are only possible when the accuracy on both the *x*-axis, that is, the certainty in morphological classification and the accuracy on the *y*-axis, that is, the inferred galaxy property is of equally good quality, thanks to the multiwavelength photometry (Wright et al. 2016). The voting in Hart et al. (2017) was high accuracy with an earlier iteration of Galaxy Zoo but the accuracy in their star formation measure did not quite match that of their Galaxy Zoo classifications, smoothing out the relation between arms and star formation rate. The combination of voting and SED accuracy allowed Porter-Temple et al. (2022) to improve on the Hart et al. (2017) result. For similar reasons, we caution against the use of Galaxy Zoo questions on morphological details (<kpc in size) for redshifts over $z = 0.1$ for either KiDS or DESI, as these correspond to more than the $0.''7$ spatial resolution.

8. Conclusions

In this paper, we directly compared two different iterations of the Galaxy Zoo morphology classification based on two different imaging surveys in the three GAMA equatorial fields. The images the classifications are based on, differ in depth, construction of RGB image, resolution, and target redshift range. The DESI-GZ catalogue is the result of ZOOBOT predictions based on all of DESI-LS trained classifications.

We found that for individual galaxies, the voting fractions can often be quite different (several tens of percent; see Figs. 5–14). However, by and large the voting between both iterations agrees, especially for the populations at large.

Reproducing the results from (Porter-Temple et al. 2022), we find the same trends as they did using the DESI-GZ catalogue.

With similar constraints on redshift, the DESI-GZ catalogue is suitable for similar work on morphological details.

We note that the DESI-GZ ZOOBOT predictions has a higher fraction of 'smooth' classifications for galaxies that have more 'disc or features' in T00. This is likely a combination of distance and depth of DESI imaging, hiding lower surface brightness features such as spiral arms and the discs of galaxies and an effect of weighting in the ZOOBOT classifications.

Data Availability. There are two catalogs that will be made public on the GAMA website.

References

Astropy Collaboration, et al. 2013, A&A, 558, A33
Astropy Collaboration, et al. 2018, AJ, 156, 123
Baldry, I. K., et al. 2018, MNRAS, 474, 3875
de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. 2013, ExA, 35, 25
de Jong, J. T. A., et al. 2015, A&A, 582, A62
de Jong, J. T. A., et al. 2017, A&A, 604, A134
Dey, A., et al. 2019, AJ, 157, 168
Driver, S. P. 2021, arXiv e-prints, arXiv:2102.12089
Driver, S. P., et al. 2009, A&G, 50, 050000
Driver, S. P., et al. 2018, MNRAS, 475, 2891
Driver, S. P., et al. 2022, MNRAS, 513, 439
Glass, D. H. W. 2024, Master's thesis, University of Central Lancashire. http://doi.org/10.17030/uclan.thesis.00042515
Hart, R. E., et al. 2017, MNRAS, 468, 1850
Holwerda, B. W., et al. 2019, AJ, 158, 103
Holwerda, B. W., et al. 2022, MNRAS, 513, 1972
Keel, W. C., et al. 2013, PASP, 125, 2
Keel, W. C., et al. 2022, arXiv e-prints, arXiv:2202.01098
Keel, W. C., et al. 2023, AJ, 165, 166
Kelvin, L. S., et al. 2014, MNRAS, 439, 1245
Kelvin, L. S., et al. 2018, MNRAS, 477, 4116
Knabel, S., et al. 2020, AJ, 160, 223
Knabel, S., et al. 2023, MNRAS, 520, 804
Kuijken, K., et al. 2019, A&A, 625, A2
Lintott, C. J., et al. 2008, MNRAS, 389, 1179
Liske, J., et al. 2015, MNRAS, 452, 2087
Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163
Masters, K. L., et al. 2011, MNRAS, 411, 2026
Moffett, A. J., et al. 2016, MNRAS, 457, 1308
Nair, P. B., & Abraham, R. G. 2010a, ApJS, 186, 427
Nair, P. B., & Abraham, R. G. 2010b, ApJ, 714, L260
Porter, L. E., et al. 2023, MNRAS, 524, 5768
Porter-Temple, R., et al. 2022, MNRAS, 515, 3875
Smethurst, R. J., et al. 2015, MNRAS, 450, 435
Smethurst, R. J., et al. 2022, MNRAS, 510, 4126
Smith, B. J., Giroux, M. L., & Struck, C. 2022a, AJ, 164, 146
Smith, D., et al. 2022b, arXiv e-prints, arXiv:2211.08355
Smith, M. J., Arora, N., Stone, C., Courteau, S., & Geach, J. E. 2021, MNRAS, 503, 96
Stone, M. B., et al. 2023, ApJ, 946, 116
Tan, M., & Le, Q. V. 2019, arXiv e-prints, arXiv:1905.11946
Walmsley, M., Ferguson, A. M. N., Mann, R. G., & Lintott, C. J. 2019, MNRAS, 483, 2968
Walmsley, M., et al. 2021a, MNRAS, 509, 3966

Walmsley, M., et al. 2021b, MNRAS, 513, 1581
Walmsley, M., et al. 2023a, MNRAS, 526, 4768
Walmsley, M., et al. 2023b, JOSS, 8, 5312
Willett, K. W., et al. 2013, MNRAS, 435, 2835
Wright, A. H., et al. 2016, MNRAS, 460, 765

Appendix A. Catalogue Descriptions

Here we describe the GAMA KiDS and DESI based catalogues to accompany this paper.

Tables A1 and A2 list the entries in the Galaxy Zoo KiDS classification catalog. These include a CATAID to identify the GAMA source and totals and fractions of voting on these objects.

Table A3 is the full listing of entries in the DESI Galaxy Zoo catalogue as described in Walmsley et al. (2023a). These the the GAMA CATAID and right ascention and declination used to match with the entries in the DESI catalogue. This catalogue contains only fractions of votes as these are predicted by zoobot.

Table A1. The KiDS Galaxy Zoo catalogue entries. Total vote numbers for each question.

Name	Unit	Description
region	–	The GAMA region
subject_id	–	The identifier for subject
survey_id	–	The survey (GAMA) ID, identical to CATAID
zooniverse_id	–	The Galaxy Zoo object identification
features_total	N_{T00}	Number of votes for T00
edgeon_total	N_{T01}	Number of votes for T01
bar_total	N_{T02}	Number of votes for T02
spiral_total	N_{T03}	Number of votes for T03
bulge_total	N_{T07}	Number of votes for T07
spiralwinding_total	N_{T05}	Number of votes for T05
spiralnumber_total	N_{T06}	Number of votes for T06
bulgeshape_total	N_{T07}	Number of votes for T07
round_total	N_{T08}	Number of votes for T08
mergers_total	N_{T09}	Number of votes for T09
oddtype_total	N_{T10}	Number of votes for T10
discuss_total	N_{T11}	Number of votes for T11
odd_total	N_{votes}	Number of votes for T00

Table A2. The KiDS Galaxy Zoo catalogue entries. Vote fraction for each question.

Name	Name (unit)	Description
CATAID	–	The CATAID to match to GAMA catalogues
features_smooth_frac	f_{smooth}	Fraction of votes T00 for "smooth" (A0)
features_features_frac	$f_{features}$	Fraction of votes T00 for "features" (A1)
features_star_or_artifact_frac	$f_{artifact}$	Fraction of votes T00 for "star/artifact" (A2)
edgeon_yes_frac	$f_{edge-on}$	Fraction of votes T01 for yes edge-on (A0)
edgeon_no_frac	$f_{notedge-on}$	Fraction of votes T01 for not edge-on (A1)
bar_bar_frac	f_{bar}	Fraction of votes T02 for yes barred (A0)
bar_no_bar_frac	f_{nobar}	Fraction of votes T02 for no bar (A1)
spiral_spiral_frac	$f_{spiralarms}$	Fraction of votes T03 for yes spiral arms (A0)
spiral_no_spiral_frac	$f_{nospiralarms}$	Fraction of votes T03 for no spiral arms (A1)
bulge_no_bulge_frac	$f_{nobulge}$	Fraction of votes T04 for "no bulge" (A0)
bulge_obvious_frac	$f_{obvious}$	Fraction of votes T04 for "obvious bulge" (A1)
bulge_dominant_frac	$f_{dominant}$	Fraction of votes T04 for "dominant bulge" (A2)
spiralwinding_tight_frac	f_{tight}	Fraction of votes T05 for "Tight spiral pattern" (A0)
spiralwinding_medium_frac	f_{medium}	Fraction of votes T05 for "Medium spiral pattern" (A1)
spiralwinding_loose_frac	f_{loose}	Fraction of votes T05 for "Loose spiral pattern" (A2)
spiralnumber_1_frac	f_{1arm}	Fraction of votes T06 for a single spiral arm (A0)
spiralnumber_2_frac	f_{2arm}	Fraction of votes T06 for two spiral arms (A1)
spiralnumber_3_frac	f_{3arm}	Fraction of votes T06 for three spiral arms (A2)
spiralnumber_4_frac	f_{4arm}	Fraction of votes T06 for four spiral arms (A3)
spiralnumber_more_than_4_frac	f_{5+arm}	Fraction of votes T06 for more than 4 spiral arms (A4)
bulgeshape_rounded_frac	f_{round}	Fraction of votes T07 for a rounded bulge (A0).
bulgeshape_boxy_frac	f_{boxy}	Fraction of votes T07 for a boxy bulge (A1)
bulgeshape_no_bulge_frac	$f_{nobulge}$	Fraction of votes T07 for no bulge (A2)
round_completely_round_frac	f_{round}	Fraction of votes T08 for round shape (A0)
round_in_between_frac	$f_{in-between}$	Fraction of votes T08 for in-between shape (A1)
round_cigar_shaped_frac	f_{cigar}	Fraction of votes T08 for cigar shape (A2)
mergers_merging_frac	$f_{merging}$	Fraction of votes T09 for merging galaxies.
mergers_tidal_debris_frac	f_{debris}	Fraction of votes T09 for tidal debris or tails
mergers_both_frac	f_{both}	Fraction of votes T09 for both merging and tidal debris
mergers_neither_frac	$f_{neither}$	Fraction of votes T09 for neither merging or tidal debris
oddtype_none_frac	f_{noodd}	Fraction of votes T10 for no odd features (X0)
oddtype_ring_frac	f_{ring}	Fraction of votes T10 for a ring feature (X1)
oddtype_lens_or_arc_frac	f_{arc}	Fraction of votes T10 for lens/arc feature (X2)
oddtype_irregular_frac	$f_{irregular}$	Fraction of votes T10 for irregular appearance (X3)
oddtype_other_frac	f_{other}	Fraction of votes T10 for "other" features (X4)
oddtype_dust_lane_frac	$f_{dustlane}$	Fraction of votes T10 for a dust lane feature (X5)
oddtype_overlapping_frac	$f_{overlap}$	Fraction of votes T10 for an overlapping pair (X6)
discuss_yes_frac	f_{votes}	Fraction of votes T11 to flag for discussion
discuss_no_frac	f_{votes}	Fraction of votes T11 to not flag for discussion

Table A3. The DESI Galaxy Zoo catalogue entries. CATAID and basic information from GAMA target catalogue and the vote fraction for each question.

Name	Name (unit)	Description
CATAID	–	GAMA catalog ID
GAMA_RA	(degree)	GAMA catalog right ascension
GAMA_DEC	(degree)	GAMA catalog declination
dr8_id	–	DESI DR8 ID
ra	(degree)	DESI catalog right ascension
dec	(degree)	DESI catalog declination
brickid	–	DESI brick number
objid	–	DESI Object ID
hdf5_loc	–	HDF5 location
smooth-or-featured_smooth_fraction	f_{smooth}	T00 fraction of votes for "smooth" (A0)
smooth-or-featured_featured-or-disk_fraction	$f_{featured}$	T00 fraction of votes for "featured" (A1)
smooth-or-featured_artifact_fraction	$f_{artifact}$	T00 fraction of votes for "artifact or star" (A2)
disk-edge-on_yes_fraction	f_{edgeon}	T01 fraction of votes for "edgeon" (A0)
disk-edge-on_no_fraction	f_{edgeon}	T01 fraction of votes for "not edgeon" (A1)
has-spiral-arms_yes_fraction	f_{spiral}	T03 fraction of votes for "spiral" (A0)
has-spiral-arms_no_fraction	$f_{no-spiral}$	T03 fraction of votes for "no spiral" (A1)
bar_strong_fraction	$f_{strongbar}$	T03 fraction of votes for "strong bar" (A0)
bar_weak_fraction	$f_{weakbar}$	T03 fraction of votes for "weak bar" (A1)
bar_no_fraction	f_{nobar}	T03 fraction of votes for "no bar" (A2)
bulge-size_dominant_fraction	$f_{dominant-bulge}$	T04 fraction of votes for "dominant bulge" (A0)
bulge-size_large_fraction	$f_{large-bulge}$	T04 fraction of votes for "large bulge" (A1)
bulge-size_moderate_fraction	$f_{moderate-bulge}$	T04 fraction of votes for "moderate bulge" (A2)
bulge-size_small_fraction	$f_{small-bulge}$	T04 fraction of votes for "small bulge" (A3)
bulge-size_none_fraction	$f_{no-bulge}$	T04 fraction of votes for "no bulge" (A4)
spiral-winding_tight_fraction	f_{tight}	T05 fraction of votes for "tightly" wound spiral arms (A0)
spiral-winding_medium_fraction	f_{medium}	T05 fraction of votes for "medium" wound spiral arms (A1)
spiral-winding_loose_fraction	f_{loose}	T05 fraction of votes for "loosely" wound spiral arms (A2)
spiral-arm-count_1_fraction	f_{1-arm}	T06 fraction of votes for 1 spiral arm (A0)
spiral-arm-count_2_fraction	f_{2-arm}	T06 fraction of votes for 2 spiral arms (A1)
spiral-arm-count_3_fraction	f_{3-arm}	T06 fraction of votes for 3 spiral arms (A2)
spiral-arm-count_4_fraction	f_{4-arm}	T06 fraction of votes for 4 spiral arms (A3)
spiral-arm-count_more-than-4_fraction	f_{5+arm}	T06 fraction of votes for 5+ spiral arms (A4)
spiral-arm-count_cant-tell_fraction	$f_{?-arm}$	T06 fraction of votes for unclear number of spiral arms (A5)
edge-on-bulge_boxy_fraction	f_{round}	T07 fraction of votes for "rounded" bulge (A0)
edge-on-bulge_none_fraction	f_{round}	T07 fraction of votes for "boxy" bulge (A1)
edge-on-bulge_rounded_fraction	f_{round}	T07 fraction of votes for "round" galaxy (A2)
how-rounded_round_fraction	f_{round}	T08 fraction of votes for "round" galaxy (A0)
how-rounded_in-between_fraction	$f_{in-between}$	T08 fraction of votes for "in-between" (A1)
how-rounded_cigar-shaped_fraction	f_{cigar}	T08 fraction of votes for "cigar" (A2)
merging_none_fraction	$f_{merging}$	T09 fraction of votes for "merging" (A0)
merging_minor-disturbance_fraction	f_{tidal}	T09 fraction of votes for "tidal debris" (A1)
merging_major-disturbance_fraction	f_{both}	T09 fraction of votes for "both" merging and tidal debris (A2)
merging_merger_fraction	$f_{neither}$	T09 fraction of votes for "neither" merging or tidal debris (A3)