Task 1.1: If the field of view is $180°$, then θ can take on values in the range $[-90°, +90°]$. This is problematic due to the $\tan\theta$ term, which grows unbounded as $|\theta| \to 90°$, which in turn causes the pixel coordinates to grow unbounded. Hence, if the field of view is $180°$, then the undistorted image would need to have infinitely high resolution, but even before then the resolution required to preserve reasonable detail across the image will quickly become impractically large. In this sense, an ultra-wide camera can be said to "use" distortion in order to preserve detail more evenly across the image.

Task 1.2:

(a) The principal point parameters c_x, c_y must be multiplied by $1/2$. The terms f_x and f_y must be scaled by $1/2$. The distortion coefficients k_i, p_i are unchanged.

(b) All parameters are unchanged except c_x, which should be adjusted to $c'_x = c_x - l$.

Task 1.3: $k_1 < 0$ for the left image and $k_1 > 0$ for the right image.

Task 1.4: $9 + 6N$.

Task 1.5: The scale of the checkerboard pattern only affects the magnitude of the translation vectors. It would be like specifying the checkerboard point coordinates in millimeters versus meters; the result is just that the translation vectors would be in millimeters or meters, while the intrinsics and rotations are unaffected.

You can see this mathematically. Let $\mathbf{X}$ be a point on the checkerboard pattern, in some coordinate frame attached to the pattern. Then, its "calibrated image coordinates" (x, y) in one image are given by

$$\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{bmatrix} = \mathbf{R}\mathbf{X} + \mathbf{t} \tag{1}$$

where $(x, y) = (\tilde{x}/\tilde{z}, \tilde{y}/\tilde{z})$, and $(\mathbf{R}, \mathbf{t})$ is the transformation from pattern coordinates to camera coordinates for this image. Now suppose every point on the checkerboard is scaled by a factor μ to $\mathbf{X}' = \mu\mathbf{X}$, and define the equivalently scaled translation vector $\mathbf{t}' = \mu\mathbf{t}$. Then, the scaled point's projection, using the scaled translation vector, is

$$\tilde{\mathbf{x}}' = \mathbf{R}\mathbf{X}' + \mathbf{t}' = \mu(\mathbf{R}\mathbf{X} + \mathbf{t}). \tag{2}$$

However, the scaling factor μ can be dropped because the left-hand side is homogeneous, and thus we obtain the same coordinates as for the unscaled points. Hence, the predicted image coordinates in all images are unchanged if we scale the translation vectors by the same factor.

Task 1.6: If only a single image is taken of the pattern, such that it appears in a small area in the center of the image, then it provides little information about distortion, which is strongest near the image edges. Then the reprojection error can be small, simply because the error isn't measured everywhere.

Task 1.7: Let $(X, Y, 0)$ be the coordinates of a point on the pattern in the pattern's coordinate frame. Since the pattern appears frontoparallel, the rotation is the identity, and its pixel coordinates are therefore

$$u = c_x + f_x(X + t_x)/t_z, \tag{3}$$
$$v = c_y + f_y(X + t_y)/t_z, \tag{4}$$

where t_x, t_y, t_z is the camera translation, which is to be estimated along with f_x, f_y, c_x, c_y. This is problematic, because we may multiply t_z by some factor and divide f_x and f_y by the same factor and get the exact same coordinates (u, v). We may also add a constant to c_x and subtract the same constant multiplied by f_x/t_z from t_x and get the same coordinates. Hence, changing the translation parameters becomes indistinguishable from changing the intrinsics.

Task 2.1: Let u_1 and u_2 be the horizontal pixel coordinate of some hypothetical 3D point, computed using the nominal value of f_x, and the nominal value plus 1.96 standard deviations, respectively. The task is asking us to find the largest possible value of the difference $|u_1 - u_2|$. As mentioned in the assignment, error in the focal length has the largest effect near the image sides or corners. To compute u_1 and u_2, we should therefore use the 3D coordinates of a point that would project to $(W, 0)$, under the nominal value of f_x. We only need the value of $(x + \delta_x)$ for this point, which we can solve for as follows:

$$\begin{aligned} u &= W \Rightarrow & (5)\\ c_x + f_x(x + \delta_x) &= W \Rightarrow & (6)\\ (x + \delta_x) &= (W - c_x)/f_x & (7)\\ &= (2816 - 1370.05852)/2359.40946 \approx 0.61. & (8) \end{aligned}$$

Then we can compute $|u_1 - u_2|$,

$$|u_1 - u_2| = 1.96 \cdot 0.84200|(x + \delta_x)| = 1.01 \text{ pixels.} \tag{9}$$

Task 2.2: The largest error in the image resulting from omitting p_1 is (most likely) at most 2.47 pixels, and for p_2 at most 16.47 pixels. One could argue for omitting p_1, as its effect is close to one pixel, but p_2 can probably not be omitted.

Comment: It may be unclear why the task asks you to involve the standard deviations in your calculations. For simplicity I will just talk about p_1. Suppose the estimated value of p_1 is exactly 0, but its standard deviation is such that p_1 is likely to be in the range [-0.5, +0.5]. If you want to analyze the importance of using the estimated value of p_1 versus leaving it at zero, then you can obviously conclude that it can be left at zero, because it already is zero. However, if you want to analyze the importance of calibrating for tangential distortion for your camera (regardless of whether it was calibrated well or not in this particular case), then you should compare the effect of p_1 being +0.5 versus 0, and p_1 being -0.5 versus 0. If either comparison is significant, then you can conclude that you do need to calibrate for tangential distortion for your camera. In this particular example, it would also suggest that you should redo the calibration (why? and how can you decide this if p_1's estimate is not just zero?).